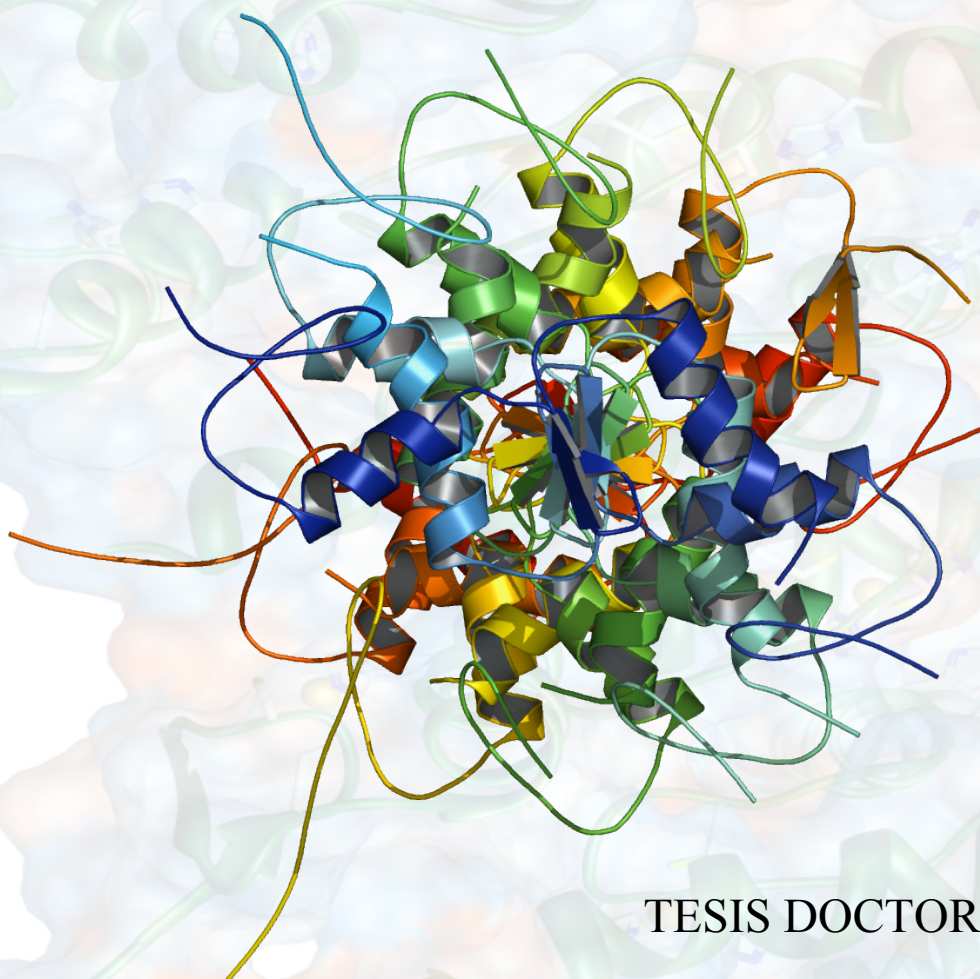


Universidad Autónoma de Madrid
Departamento de Biología Molecular
Facultad de Ciencias

*Estudios estructurales de BRMS1
y Tax1BP1, proteínas
involucradas en la vía de
activación de NF κ B*



TESIS DOCTORAL

Mercedes Spínola Amilibia
Madrid, 2009

Universidad Autónoma de Madrid

Departamento de Biología Molecular

Facultad de Ciencias

*Estudios estructurales de
BRMS1 y Tax1BP1, proteínas
involucradas en la vía de
activación de NFκB*

*Memoria presentada por
Mercedes Spínola Amilibia para
optar al grado de Doctora en
Ciencias Biológicas*

Madrid, Enero de 2009

El trabajo recogido en esta memoria ha sido realizado en el Centro Nacional de Investigaciones Oncológicas (CNIO), en Madrid, bajo la dirección de los Drs. Jerónimo Bravo Sicilia y José Rivera Torres.



Dr. Jerónimo Bravo Sicilia

y

Dr. José Rivera Torres

CERTIFICAN

Que el presente trabajo de investigación, titulado “Estudios estructurales de *BRMS1* y *Tax1BP1*, proteínas involucradas en la vía de activación de NFκB”, ha sido realizado por D^a. Mercedes Spínola Amilibia, bajo nuestra dirección en el Centro Nacional de Investigaciones Oncológicas de Madrid y que, salvo mejor criterio del tribunal que ha de juzgarla, reúne todas las condiciones exigidas para optar al grado de Doctora en Ciencias Biológicas.

Y para que así conste, expedimos el presente certificado en Madrid a 13 de Enero de 2009.

Fdo.: Dr. Jerónimo Bravo Sicilia

Fdo.: Dr. José Rivera Torres

A Diego

Resumen/Summary	17
Introducción	21
1. Metástasis	23
2. Genes Supresores de Metástasis	24
2.1 BRMS1	28
2.1.1 Características estructurales de BRMS1; a nivel genómico y proteico	28
2.1.2 Mecanismos moleculares de BRMS1	30
3. Tax1BP1	33
3.1 Características estructurales de Tax1BP1	33
3.2 Ligandos celulares de Tax1BP1 y sus implicaciones en procesos de ubiquitinación	35
3.2.1 Tax1	35
3.2.2 A20	35
3.2.3 TRAF6	36
3.2.4 RIP1	37
3.2.5 Itch (ligasa E3)	37
4. BRMS1 y Tax1BP1, reguladores negativos de la activación de NFκB	37
4.1 Aspectos moleculares de la ruta de activación de NFκB	38
4.2 BRMS1 como modulador de la ruta de activación de NFκB	40
4.2.1 Modificaciones postraduccionales sobre el factor de transcripción NFκB por parte de BRMS1	41
4.3 Relevancia de Tax1BP1 en la ruta de activación de NFκB	42
4.3.1 Procesos de ubiquitinación en regulación de la ruta NFκB	42
Objetivos	45
Materiales y Métodos	49
1. Clonaje de los diferentes amplicones en vectores de expresión procariota	51
1.1 Generación de amplicones y construcción de plásmidos recombinantes	51
1.2 Electroforesis de ADN	53
2. Expresión de proteínas recombinantes	55
2.1 Expresión de proteínas derivadas con selenio-metionina (SeMet - BRMS151-98)	56

2.2	<i>Expresión de proteínas marcadas con Nitrógeno 15 (15N-BRMS151-84)</i>	56
3.	<i>Purificación de proteínas</i>	57
3.1	<i>Purificación de las construcciones con cola de histidinas</i>	58
3.2	<i>Construcciones fusionadas a GST</i>	60
4.	<i>Obtención de ubiquitina y poli4-Ubiquitina humanas</i>	62
5.	<i>Análisis físico-químicos y cuantificación de proteínas</i>	62
5.1	<i>Cuantificación de proteínas</i>	62
5.1.1	<i>Espectroscopía de luz ultravioleta a 280 nm.</i>	62
5.1.2	<i>Método de colorimetría: Bradford</i>	63
5.2	<i>Análisis de los parámetros físico-químicos de cada proteína</i>	63
6.	<i>Ensayos biofísicos</i>	63
6.1	<i>Espectroscopía de Dicroísmo Circular (CD)</i>	63
6.2	<i>Dispersión dinámica de la luz (DLS)</i>	64
7.	<i>Ensayos de caracterización de interacción proteína-proteína</i>	64
7.1	<i>Calorimetría de titulación isotérmica (ITC)</i>	64
7.2	<i>Resonancia Magnética Nuclear (RMN)</i>	64
7.3	<i>Cromatografía de exclusión por tamaño molecular</i>	65
8.	<i>Ensayos de cristalización</i>	65
8.1	<i>Cristalización inicial</i>	66
8.2	<i>Optimización de cristales</i>	66
9.	<i>Montaje de cristales y vitrificación</i>	68
10.	<i>Resolución de estructuras</i>	69
10.1	<i>Recolección de datos de difracción</i>	69
10.2	<i>Reducción de datos</i>	70
10.3	<i>Resolución de fases experimentales</i>	70
10.3.1	<i>Obtención de cristal derivado de átomo pesado</i>	71
10.3.2	<i>Métodos de resolución de estructura usando señal anómala</i>	72
10.4	<i>Construcción del modelo</i>	73
10.5	<i>Afinado de la estructura</i>	73
10.6	<i>Validación de la estructura</i>	73
	Resultados	75
	BRMS1.	79

1. Diseño de las construcciones objeto de estudio	81
2. Resultados generales de las construcciones clonadas	83
3. Expresión, purificación, caracterización y cristalización del primer motivo “coiled-coil” de BRMS1 (BRMS151-84) (Spínola-Amilibia et al., 2008)	85
3.1 Expresión	85
3.2 Purificación	87
3.2.1 Cromatografía de afinidad	87
3.2.2 Proteólisis y eliminación de la cola de histidinas por columna de afinidad	88
3.2.3 Cromatografía de exclusión por tamaño molecular	89
3.3 Cristalización de BRMS151-84	91
3.3.1 Cristalización inicial	91
3.3.2 Optimización de cristales	92
3.4 Resolución de la estructura de BRMS151-84	93
3.4.1 Recolección de datos de difracción del cristal nativo de BRMS151-84	93
3.4.2 Recolección de datos de difracción del cristal derivado de mercurio de BRMS151-84	94
3.4.3 Procesado de datos y estructura de BRMS151-84	98
3.4.3.1 Monster, análisis de interacciones entre las dos cadenas de los motivos “coiled-coil”	105
3.4.4 Validación de la estructura de BRMS151-84	107
4. Expresión, purificación y cristalización de BRMS1 aa51-98.	109
5. Expresión y purificación de NMI aa16-64, proteína de interacción con BRMS1.	116
5.1 Expresión	116
5.2 Purificación	117
6. Caracterización de las interacciones proteína-proteína: Evidencias en la formación de complejos.	121
6.1 Co-expresión y co-purificación del complejo BRMS1 FL + NMI FL (secuencias completas)	121
6.2 Caracterización de interacción BRMS1 aa51-84 - NMI aa16-64	123
6.2.1 Espectrometría de resonancia magnética nuclear (RMN)	123
6.2.2 Calorimetría de titulación isotérmica (ITC)	126
7. Ensayos de cristalización de los complejos caracterizados biofísicamente; BRMS1FL + NMIFL y BRMS151-84 + NMI16-64	127
Tax1BP1.	129
1. Diseño de construcciones de la proteína Tax1BP1	131

2.	<i>Expresión, purificación, cristalización y resolución de la estructura de dominios de unión a Ubiquitina (UBZs) de Tax1BP1</i>	132
2.1	<i>Purificación de UBZ1+2 (aa 725-789)</i>	133
2.1.1	<i>Cromatografía de afinidad de His-PP-UBZ1+2</i>	133
2.1.2	<i>Proteolisis y cromatografía de exclusión por tamaño molecular</i>	134
2.2	<i>Estructura de UBZ1+2</i>	137
3.	<i>Evidencias de interacciones de UBZ1+2 de Tax1BP1 con ubiquitina, poli-4Ub e Itch, por co-elución en una columna de exclusión de tamaño molecular.</i>	146
3.1	<i>Interacción entre UBZ1+2 y ubiquitina</i>	147
3.2	<i>Interacción entre UBZ1+2 e Itch</i>	150
4.	<i>Ensayos de cristalización de UBZ1+2 en complejo con Ub, poli-4Ub e ItchWW3+WW4</i>	153
	Discusión	155
1.	<i>Estudios estructurales de BRMS1 y su interacción con NMI</i>	158
1.1	<i>Caracterización de la interacción entre BRMS1 y NMI</i>	159
1.2	<i>Estructura del primer motivo “coiled-coil” de BRMS1</i>	160
1.2.1	<i>Twister, análisis de motivos “coiled-coil”</i>	162
2.	<i>Tax1BP1 y su papel en mecanismos de ubiquitinación</i>	168
2.1	<i>Relevancia biológica de los dominios UBZ de Tax1BP1</i>	169
2.2	<i>Estudios estructurales de UBZ1+2</i>	169
2.3	<i>UBZ, como dominio de unión a ubiquitina</i>	171
2.4	<i>Interacción de UBZ1+2 con la ligasa de ubiquitina Itch, a través de las secuencias consenso PPxY</i>	176
	Conclusiones	179
	Anexo	183
1.	<i>Resolución de estructuras mediante Cristalografía de Rayos X</i>	185
1.1	<i>Difracción de Rayos X</i>	186
1.2	<i>Reducción de datos</i>	188
1.3	<i>Resolución de fases experimentales</i>	189
1.4	<i>Construcción y afinado del modelo</i>	191
1.5	<i>Validación de la estructura</i>	192

	<i>Índice</i>
<i>Bibliografía</i>	<i>195</i>
<i>Referencias</i>	<i>213</i>
<i>Tablas y Figuras</i>	<i>219</i>
<i>Abreviaturas</i>	<i>227</i>

Resumen/Summary

Metastasis is a life-threatening complication of cancer that represents 90% of cancer-related deaths. It is the process by which tumor cells disseminate from the primary tumor, migrate through the basement membrane, survive in the circulatory system, invade into a secondary site, and start to proliferate (Gupta y Massagué, 2006; Pantel y Brakenhoff, 2004; Steeg, 2003; Steeg y Theodorescu, 2008). This process is mainly developed by the de-regulation of vital pathways controlling the biological mechanism.

One of them is the signalling pathway that mediate the activation of the NFκB transcription factor, a central coordinator of the immune response. It has become clear that NFκB signalling has a critical role in cancer development and progression. It also regulate tumour angiogenesis and invasiveness (Karin, 2006; Lin y Karin, 2003), being correlated with up-activation of the transcription factor level. Nowadays, the mayor goal has been to determine the molecular networks, looking at genes that induce or inhibit metastasis. For this reason, we focused on two negative regulators of NFκB pathway, which are, BRMS1 and Tax1BP1.

Breast cancer metastasis suppressor 1 (BRMS1) is a member of a functional family of genes known as metastasis suppressors that have the ability to inhibit the appearance of macroscopic metastases in breast, melanoma and ovarian tumours without affecting the growth of the primary tumour, by acting at different steps of the metastatic cascade (Phadke *et al.*, 2008). Very little is known about the molecular aspects of the antimetastatic properties of BRMS1, and therefore the structural characterization of BRMS1 can provide important information on the molecular mechanisms involved in this protein function.

In the present study, we show the tridimensional structure of the first predicted coiled-coil motif of human BRMS1. In addition, it has been previously determined the participation of the N-terminal coiled-coil of this protein in interactions with some relevant proteins to cancer (Rivera *et al.*, unpublished results). One of them is a co-factor of transcription, named NMI, N-Myc and STAT interactor protein. The rescued NMI fragment in a yeast to hibrid assay included the predicted coiled-coil motif. It is reported that coiled-coil regions are well known protein–protein interaction modules (Lupas, 1996), and for this reason, we selected this motif for structural studies of the complex. NMR and ITC experiments shown that the two coiled-coil motif interact upon tested conditions.

On the other hand, it is reported that ubiquitination regulates at least three steps in the NFκB pathway (Chen, 2005; Haglund y Dikic, 2005; Hoeller *et al.*, 2006). Tax1BP1 or “Tax Binding Protein 1” is one of the proteins implicated in this process. The interaction between this protein and ubiquitin is through its carboxyl-terminal end, which contains two zinc finger domains (*UBZ*) (Iha *et al.*, 2008). We obtain the tridimensional conformation of both domains, and we concluded that only *UBZ2* appears to be a ubiquitin binding C₂H₂ zinc finger domain. The consensus sequence established in the UBD (Ubiquitin Binding Domains) is not conserved amongst *UBZ1* and *UBZ2*, which suggests that the interaction with ubiquitin should take place in a different manner.

TAX1BP1 seems to act in NFκB signalling as an essential adaptor between ubiquitin-editing enzyme A20 and its targets. It is demonstrated that the molecule TAX1BP1 recruits the E3 ubiquitin ligase Itch to A20 via two *PPXY* motifs (Shembade *et al.*, 2008). We had evidences that this interaction takes place through *PPxY* of Tax1BP1 protein and tryptophan rich region domains 3 and 4 (*WW*₃ and *WW*₄) located within Itch.

The two structures solved by X Ray Crystallography in this work, represent the first structural reports of BRMS1 and Tax1BP1. The knowledge of the tridimensional conformation of both domains, can be used for understanding the biological mechanisms involving BRMS1 and Tax1BP1, as well as their role in the NFκB activation, providing attractive targets for new chemopreventive and chemotherapeutic approaches.

Introducción

El cáncer representa la tercera causa de muerte a nivel mundial, tras las enfermedades cardiovasculares y las infecciosas. En España, representa la primera causa de mortalidad entre los hombres y la segunda entre las mujeres, tras las enfermedades cardiovasculares, registrándose en el año 2005, 96.499 defunciones, que representa el 25 % de muertes totales, según datos del Centro Nacional de Epidemiología.

Mientras la posibilidad de tratar con éxito tumores bien localizados ha mejorado ostensiblemente en los últimos años, la metástasis representa un reto terapéutico continuo y se mantiene como la causa más común de muerte relacionada con el cáncer.

1. Metástasis

La metástasis es un proceso que consiste en la formación progresiva de focos secundarios de crecimiento celular tumoral procedente de una lesión o foco primario. Este proceso engloba varias etapas (Figura 1), empezando por la adquisición de capacidad invasiva de las células cancerígenas primarias, que implica cambios en adhesión celular. Posteriormente se inicia una motilidad celular y proteólisis de la matriz extracelular, culminando con la intravasación de células en el torrente circulatorio directamente, o por vía linfática (Kauffman *et al.*, 2003; Langley y Fidler, 2007; Pantel y Brakenhoff, 2004; Steeg, 2003; Steeg y Theodorescu, 2008). Algunas de estas células sobreviven y son transportadas hacia otros órganos donde se extravasan y asientan, originando focos secundarios o micrometástasis, que pueden permanecer ocultas durante un cierto tiempo. Finalmente tiene lugar una progresiva colonización metastásica, lo que conlleva crecimiento y proliferación celular. Dependiendo del grado y eficacia de esta invasión y colonización, se habla entonces de capacidad metastásica celular o agresividad tumoral.

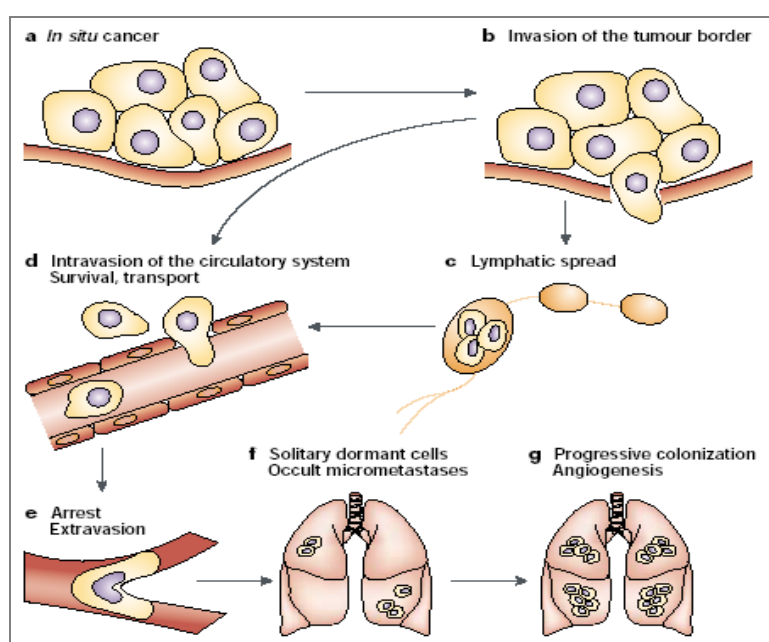


Figura 1.

Etapas en el proceso de metástasis:

a | Células cancerígenas *in situ* en tumor primario.

b | La invasión celular requiere cambios reversibles en adhesión célula-célula y célula-matriz extracelular.

Las células metastásicas pueden **c** | entrar vía linfática, o **d** | directamente a la circulación.

e | Supervivencia y detención de células tumorales con consecuente extravasación.

f | Colonización metastásica en focos distantes de células individuales, que pueden permanecer “dormidos” durante años (micrometástasis).

g | Progresivo crecimiento y metástasis angiogénica.

Figura tomada de la revisión: Steeg, *Nature Reviews Cancer* 3 (2003) 55-63 (Steeg, 2003)

En la mayor parte de los tumores, la presencia de metástasis clínico-perceptibles conlleva desgraciadamente la etiqueta de “enfermedad difícil de curar”. Se ha avanzado mucho en la correcta detección del cáncer y en la aplicación de tratamientos para eliminar tumores primarios con éxito, como son la cirugía médica y otras técnicas como aplicación de radiación. Sin embargo, estas modalidades han demostrado tener eficacia limitada contra el cáncer o sus consecuencias, ya que el verdadero problema reside en el control de la metástasis y no en el tumor primario, que con un diagnóstico precoz puede ser combatido hoy en día con cierta facilidad.

Durante los últimos 30 años, la investigación en metástasis ha experimentado grandes avances a nivel básico y clínico que aclaran muchas de las alteraciones genéticas y moleculares de esta patología (Berger *et al.*, 2004; Eccles y Welch, 2007; Langley y Fidler, 2007; Nguyen y Massagué, 2007; Pantel y Brakenhoff, 2004; Steeg y Theodorescu, 2008; Welch *et al.*, 2000). Desafortunadamente, tal conocimiento ha dado lugar a pocos avances terapéuticos en el control de la misma.

La investigación llevada a cabo durante la última década en varias disciplinas, ha dado a conocer nuevas teorías de cómo unas células del tumor primario se convierten en metastásicas. Hoy en día, existen dos corrientes de opinión:

1. Defiende que las células que originan metástasis se presentan como un grupo de variantes metastásicas, que no son genéticamente distintas a las del tumor primario, las cuales, a priori, no tendrían predisposición alguna a desarrollar metástasis. Los genes que conducen al desarrollo de metástasis son los mismos genes que llevaron a la célula a desembocar en tumor primario, tales como los oncogenes Ras y Myc, y simplemente se sucede una desregulación de los mismos (Bernards y Weinberg, 2002; Hanahan y Weinberg, 2000; van 't Veer *et al.*, 2002).
2. Explica que las células que van a adquirir potencial metastásico, lo hacen porque llevan una firma intrínseca basada en un patrón de expresión diferencial a los que rigen el tumor primario y que les confiere una predisposición metastásica (Kang *et al.*, 2003; Minn *et al.*, 2005; Smid *et al.*, 2006). En este entorno se enmarca la clasificación reciente de un grupo de genes conocidos como **genes supresores de metástasis** (Berger *et al.*, 2004; Kauffman *et al.*, 2003; Rinker-Schaeffer *et al.*, 2006; Shevde y Welch, 2003; Stafford *et al.*, 2008; Steeg, 2003), con una función común en supresión de metástasis, y ninguno de los cuales parecen estar involucrados en el crecimiento o desarrollo del tumor primario.

2. Genes Supresores de Metástasis

Estos genes, o sus proteínas codificadas, se definen por su capacidad de inhibir metástasis *in vivo*, sin afectar al crecimiento del tumor primario. Se postula que estos genes desempeñan su papel en los estadios finales de la cascada metastásica, afectando principalmente a la colonización metastásica de las células que previamente se han diseminado del tumor primario (Berger *et al.*, 2004; Kauffman *et al.*, 2003; Rinker-Schaeffer *et al.*, 2006; Shevde y Welch, 2003; Stafford *et al.*, 2008; Steeg, 2003), ya que los experimentos fueron realizados en modelo murino, donde se inocularon células

transfectadas con supresores de metástasis a través del torrente circulatorio, para estudiar únicamente el asentamiento de las mismas en focos distantes (Kauffman *et al.*, 2003). Es precisamente en los estadios finales de la metástasis donde se supone que cobra mayor relevancia profundizar en el conocimiento de los mecanismos moleculares que acontecen, con el fin de aportar información para el desarrollo futuro de dianas terapéuticas, puesto que en el momento de la detección del tumor primario ya pueden haberse producido metástasis inapreciables.

Se ha demostrado que durante la progresión clínica de cáncer a metástasis, se produce la pérdida de expresión para la mayoría de estos genes, ya que en muestras clínicas de pacientes, sus niveles de expresión muestran una relación inversa con el desarrollo de la enfermedad (Kauffman *et al.*, 2003). Este hecho permite la utilización de estos genes como marcadores independientes tanto de la enfermedad como de su progresión, además de como indicadores de la agresividad clínica (Heimann *et al.*, 1998; Hicks *et al.*, 2006; Schindl *et al.*, 2001; Seraj *et al.*, 2000a). La eliminación de solamente un paso en la cadena frustraría todo el proceso, de la misma manera, la restauración de la expresión de genes supresores de metástasis en células tumorales metastásicamente competentes, puede interrumpir un paso en la cascada metastásica, produciendo una ventaja clínica en esos pacientes en quienes el proceso metastásico no ha finalizado (Steeg, 2003).

En los últimos años, y tras la descripción en 1988 del primer supresor de metástasis NM23 (Steeg *et al.*, 1988), se han ido identificando de manera progresiva un número creciente de genes de este grupo, hasta comprender mas de 20 en la actualidad (Tabla I).

Se desconoce por el momento el mecanismo por el cual estos genes suprimen metástasis. Sin embargo, estudios bioquímicos recientes han demostrado que algunas de estas proteínas participan en vías de transducción altamente conservadas, que median respuestas celulares a factores de crecimiento, citoquinas y estrés celular, tales como las de receptores tirosín-quinasas, receptores acoplados a proteína-G y quinastas activadas por mitógenos (Berger *et al.*, 2004; Shevde y Welch, 2003; Steeg, 2003). De esta forma, actuarían a distintos niveles, como reprimiendo oncogenes, inhibiendo diferentes metaloproteinasas de la matriz extracelular, bloqueando vías de señalización (como la de NFκB), etc (Eccles y Welch, 2007; Rinker-Schaeffer *et al.*, 2006; Stafford *et al.*, 2008). Por ello, se ha visto que desempeñan papeles comunes en el control del crecimiento, en la adhesión celular o en procesos de reorganización del citoesqueleto, sugiriendo un mecanismo común de supresión de metástasis.

En la Tabla I se muestra un resumen de todas las proteínas supresoras de metástasis descritas hasta la actualidad por diferentes autores, así como los procesos biológicos en los que están involucradas y las etapas metastásicas donde se ha visto que ejercen su efecto.

Supresor	Localización celular	Función	Etapas metastásicas inhibidas
BRMS1	N (predominante) C (escaso)	- Forma parte del complejo mSin3:HDAC - Remodelaje de la cromatina - Regulación transcripcional - Reduce la señalización de fosfoinosítoles - Restablece las uniones Gap intercelulares - Inhibe la activación de NFκB	- Supervivencia en transporte - Colonización
Caderinas (-E, -N y -11)	M	- Adhesión entre célula-célula o entre célula-matriz extracelular - Cohesión homotípica	- EMT (Transporte a la matriz extracelular) - Invasión
Caspasa 8	C	- Induce apoptosis / "anoikis"	- Supervivencia en transporte
CD44	M	- Adhesión entre célula-célula o entre célula-matriz extracelular - Unión a ácido hialurónico y a osteopontina	- Migración
DCC	C	- Arquitectura citoplasmática - Señalización de las MAPK	- Invasión - Motilidad
DLC1	C	- Señalización vía Rho-GTPasa - Arquitectura del citoesqueleto	- Invasión - Migración
DRG1	C (predominante) N (escaso)	- Desconocida	
Gelsolina	C	- Molécula reguladora de actina - Arquitectura del citoesqueleto	- Migración
KA11	M	- Unión a DARC - Interacción con integrinas - Desensibilización del EGFR	- Intravasación - Supervivencia en el transporte
KISS1	C, S	- Ligando para receptor acoplado a proteína G - Inhibe la activación de NFκB	- Colonización
MKK4 MKK7 p38	C	- Señalización de la ruta de las MAPK activada por estrés	- Colonización - Migración
NM23	C, N	- Inhibición de la señalización de Ras - Regulación de histidina quinasa - Quinasa NDP	- Colonización - Migración
OGR1	M	- Señalización de receptor acoplado a proteína G	- Colonización - Migración
RhoGDI2	C	- Arquitectura del citoesqueleto - Señalización de endotelina y neuromedina U	- Colonización - Migración
RKIP	C	- Señalización Raf-MEK - Organización del citoesqueleto	- Migración - Invasión
RRM1	C	- Señalización PI3K de adhesiones focales - Arquitectura del citoesqueleto - Adhesión	- Motilidad - Invasión
SSECKs	C	- Adaptador para proteínas quinasas - Señalización de Src - Señalización de PKC - Señalización de Rho - Secreción de VEGF	- Migración - Angiogénesis
TIMPs	C, S, M	- Inhibición de metaloproteinasas - Señalización	- Migración - Angiogénesis - Invasión - Supervivencia en transporte - Colonización
CRSP3	N	- Regulación transcripcional - Unión a elementos de unión a promotor - Cofactor de la actividad SP1	- Motilidad - Colonización - Invasión
TXNIP	C (predominante) N (escaso)	- Comunicación celular - Transducción de señales - Mediador de estrés oxidativo - Regulación transcripcional	- Motilidad - Colonización - Invasión

Tabla I. Recopilación de proteínas supresoras de metástasis y proceso celular en el que están involucradas, propuestos por diferentes autores. También se muestran las etapas metastásicas afectadas por las mismas. Las letras N, C, M y S aluden a localización nuclear, citoplasmática, de membrana y secretada respectivamente. La tabla ha sido adaptada y modificada de la revisión Stafford *et al.* (Stafford *et al.*, 2008).

Aunque los mecanismos a través de los cuales estos genes suprimen metástasis tengan una misma base, puede haber diferencias en la complejidad de cómo ejecutan esta acción. Algunos de los genes supresores de metástasis descritos hasta el momento se han identificado en varios tipos celulares con una función general (Berger *et al.*, 2004; Stafford *et al.*, 2008; Steeg, 2003). En la Figura 2 se puede comprobar cómo determinadas rutas donde actúan oncogenes (Src, Rac y Raf-1) se ven afectadas a diferentes niveles por algunos miembros de esta familia de genes supresores de metástasis (Kauffman *et al.*, 2003) (Figura 2).

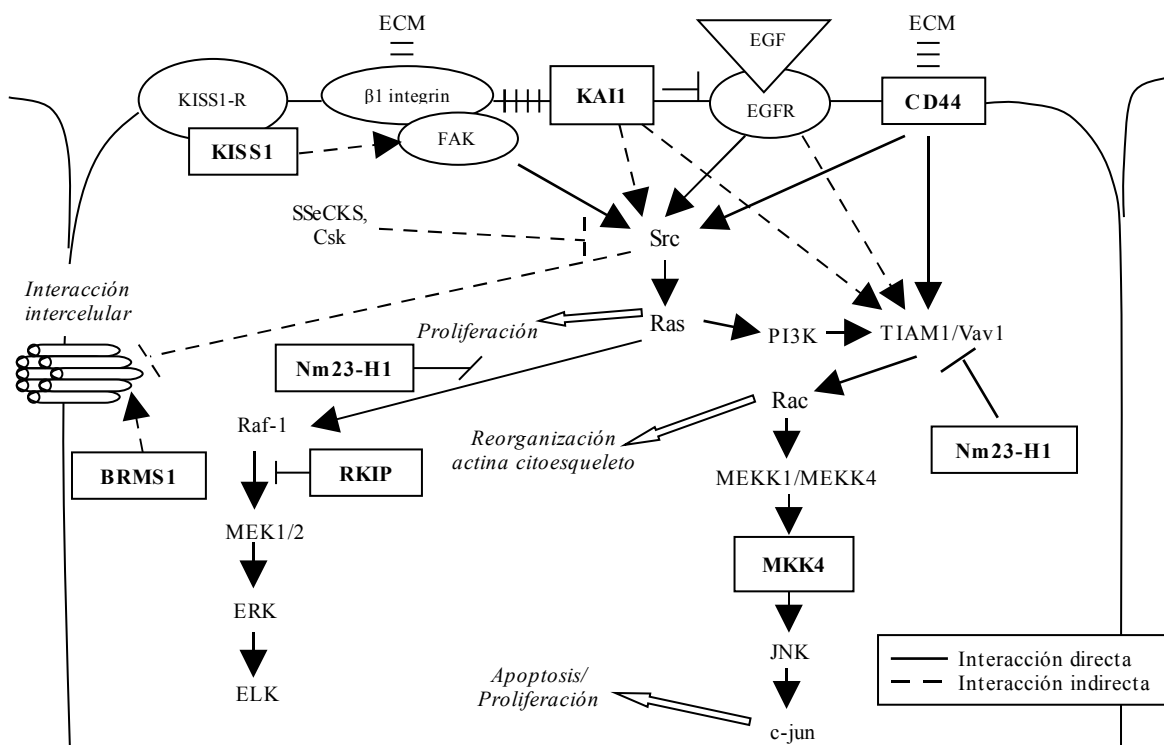


Figura 2. Modelo hipotético de interferencia entre productos de genes supresores de metástasis en vías de señalización Rac, Src y Raf-1. La figura ha sido adaptada de la revisión Kauffman *et al.* (Kauffman *et al.*, 2003).

La validación de la función de supresores de metástasis está basada en resultados obtenidos en experimentos realizados utilizando un modelo murino. En estos, el gen supresor de metástasis a estudiar es sobre-expresado en una línea celular con fenotipo altamente metastático, donde sus niveles se encuentran reducidos. Tras su inoculación en el sistema circulatorio, se analizan el desarrollo y el número de focos tumorales secundarios aparecidos en pulmones. Estos ensayos demostraron cómo en aquellos animales que fueron inoculados con células transfectadas con el gen de interés, el número de metástasis observables fue drásticamente inferior, e incluso ausentes, que las desarrolladas por los animales inoculados con las células de fenotipo metastático transfectadas con el vehículo vacío como control (Kauffman *et al.*, 2003).

2.1 BRMS1

Uno de los genes supresores de metástasis es *brms1* (en inglés “*BReast cancer Metastasis Suppressor 1*”), localizado en el brazo largo del cromosoma 11 (Phillips *et al.*, 1996; Seraj *et al.*, 2000b). Amplificaciones no aleatorias y deleciones en los brazos cortos y largos del cromosoma 11 se asocian con progresión y metástasis en pacientes con cáncer de mama (Gudmundsson *et al.*, 1995). Por ello se postuló la existencia en este cromosoma de uno o mas genes supresores de metástasis. Para su comprobación, se introdujo en células de carcinoma de mama (MDA-MB-435), una copia del cromosoma 11 intacto unido a neomicina, mediante transferencia cromosómica, tras lo que resultó un claro descenso en la formación de metástasis en ratón atímico portando las células híbridas (neo11/435). Con ello se verificó la hipótesis de la existencia de uno o mas genes supresores de metástasis en el cromosoma 11 (Phillips *et al.*, 1996).

Ensayos *in vivo*, realizados con posterioridad, demostraron su capacidad de inhibición de metástasis en modelo murino, en los que tras la inoculación de células metastásicas de mama transfectadas con el gen de interés, tanto la incidencia, como el número y tamaño de metástasis a pulmón se veían significativamente inhibidas. La expresión del gen era todavía detectable en tumores primarios, por lo cual se pudo concluir que esta proteína no inhibía la existencia ni el crecimiento del tumor primario (Seraj *et al.*, 2000b).

BRMS1 fue el primer marcador identificado en cáncer de mama y melanoma, que regula negativamente la formación de metástasis hacia pulmón y nódulos linfáticos (Seraj *et al.*, 2000b; Shevde *et al.*, 2002). Este gen se encuentra altamente conservado entre diversas especies, existiendo ortólogos de humano en especies como ratón, chimpancé, perro, toro, rata, pez cebra y mosca de la fruta, con una elevada homología en la secuencia.

2.1.1 Características estructurales de BRMS1; a nivel genómico y proteico

La estructura genómica de *brms1* consta de 9 intrones y 10 exones, el primero de los cuales no se traduce, abarcando un tamaño de aproximadamente 10 kpb. La región del extremo 5' del exón I de *brms1* contiene varios elementos reguladores que incluyen GATA-1, CREB, GATA-2 y CdxA. Sin embargo, no se ha descrito la existencia de caja TATA, lo que sugiere que se transcribe de manera independiente de ésta (Seraj *et al.*, 2000b). Se ha descrito una estructura idéntica para el homólogo murino de BRMS1 (Samant *et al.*, 2002).

Recientemente se ha descrito que el promotor de *brms1* contiene una región (-3477 a -2214) rica en GC (islas CpG), que están hipermetiladas en líneas celulares de cáncer de mama, lo que sugiere que su expresión está controlada a nivel transcripcional, mediante procesos de silenciamiento epigenético. Una posible causa del desarrollo de metástasis puede relacionarse con metilaciones aberrantes en el promotor de *brms1* y consecuente pérdida de su expresión (Metge *et al.*, 2008).

Dicho gen codifica para una proteína de 246 aminoácidos, y un peso molecular de 28461 Da, cuya localización celular es principalmente nuclear (Samant *et al.*, 2000).

La predicción de estructura secundaria (Figura 3), obtenida a través de diferentes bases de datos y herramientas de predicción informáticas (referencias 1-11), revela que la proteína BRMS1 contiene varios dominios imperfectos ricos en leucinas. Más detalladamente, estos motivos (L-x₆-L-x₆-L-x₆-L) se localizan en los rangos aminoacídicos 67-88, 131-152, 138-159, 153-174 y 160-181. A ellos se suman dos motivos con capacidad de formar bucles de cadenas α-hélice enrolladas entre sí (en inglés “coiled-coil”, término que emplearemos a partir de ahora para referirnos a estos motivos), uno desde el aminoácido 51 al 81 y el otro desde el 147 al 180. Tanto las regiones ricas en leucinas, como las mencionadas helicoidales, son dominios implicados en interacción proteína-proteína (Lupas, 1996; Mason y Arndt, 2004), lo que sugiere que la importancia funcional de esta proteína reside en estos motivos.

Se ha descrito una posible secuencia de retención en el retículo endoplasmático, situada en el extremo carboxilo terminal (C-terminal), entre los residuos 243-246, al igual que dos posibles secuencias responsables de su localización nuclear (NLS), una abarcando los aminoácidos 198-205 y la otra 239-245 (Seraj *et al.*, 2000b). Aunque los estudios que validan la función de estas secuencias como NLS todavía no han sido publicados, diversos autores han sugerido la funcionalidad de estos segmentos mediante ensayos de fraccionamiento subcelular e inmunofluorescencia y mediante análisis de delecciones (Meehan y Welch, 2003; Seraj *et al.*, 2000b).

Usando la herramienta de predicción *PROSITE*², se pudieron localizar diferentes sitios de fosforilación, estos son, para cAMP/cGMP ([R/K]₂-x-[S/T]; aminoácidos 55-58 y 240-243), para proteína Quinasa-C ([S/T]-[R/K]; aminoácidos 111-113, 147-149, 190-192 y 200-202) y para Caseína quinasa II ([S/T]-x₂-[D/E]; aminoácidos 19-22, 30-33, 37-40, 39-42, 41-44 y 46-49). Sin embargo, aún no se ha demostrado el estado de fosforilación de la proteína en condiciones experimentales.

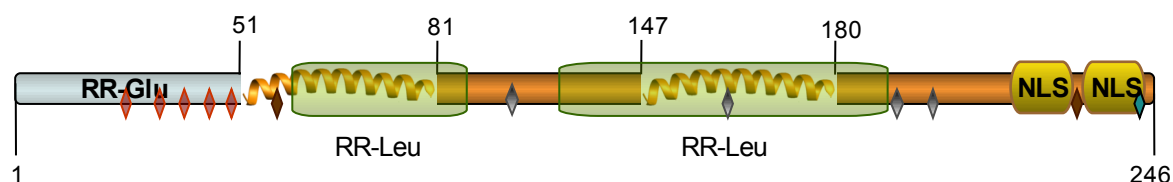


Figura 3. Representación esquemática de la predicción de motivos de la proteína BRMS1 humana. Figura realizada a partir de información extraída de diferentes herramientas de predicción (Referencias 1-11). Se muestran en números los aminoácidos inicial y final y aquellos que delimitan los motivos “coiled-coils”, aminoácido 51 a 81 y 147 a 180. En recuadros azul, verdes y amarillos se exponen regiones ricas en Glutámicos (RR-Glu), regiones imperfectas ricas en Leucinas (RR-Leu) y secuencias de localización nuclear (NLS) respectivamente. Los rombos corresponden a sitios de fosforilación: rojos de Caseína quinasa II (CK), marrones de cAMP/cGMP y grises de la proteína quinasa C (PKC). El rombo con color turquesa representa la secuencia de retención en el retículo endoplasmático.

La secuencia aminoacídica de BRMS1 contiene tres residuos cisteína (posiciones 60, 126 y 136) que aparentemente no forman parte de ningún puente disulfuro intra- o inter-proteico, puesto que la movilidad en geles de poliacrilamida no se ve afectada por el uso de agentes reductores (Seraj *et al.*, 2000b).

2.1.2 Mecanismos moleculares de BRMS1

BRMS1 fue el primer marcador identificado en cáncer de mama y melanoma, regulando negativamente la formación de metástasis hacia pulmón y nódulos linfáticos (Seraj *et al.*, 2000b; Shevde *et al.*, 2002), pero progresivamente se ha ido asociando con la misma función en diferentes tipos de cáncer, como metástasis de carcinoma de ovario humano (Zhang *et al.*, 2006) o de cáncer de vejiga (Seraj *et al.*, 2000a). Algunos artículos de reciente publicación exponen la función antimetastásica de BRMS1 en un número creciente de órganos, incrementándose su importancia en órganos tales como hueso, cerebro, riñón, páncreas y glándula adrenal (Phadke *et al.*, 2008), además de hígado, pulmón y tejido óseo, donde también se ha identificado su expresión (Hedley *et al.*, 2008). Por ello, la presencia de BRMS1 es ubicua en todo el organismo.

Los mecanismos moleculares subyacentes del proceso de supresión de metástasis mediados por BRMS1 están estrechamente relacionados con procesos de gran relevancia biológica, tales como:

- Inhibición de la ruta alternativa de señalización de supervivencia celular, **Fosfatidilinositol 3-kinasa/Akt (Protein kinase B)**. Se ha comprobado que la expresión de BRMS1 afecta a la señalización mediada por fosfoinositoles en células de cáncer de mama MDA-MB-435 (DeWald *et al.*, 2005).
- **Restauración de las uniones intercelulares tipo Gap**, involucradas en comunicación intercelular. La pérdida de las comunicaciones intercelulares “Gap” está correlacionada con progresión de metástasis (Ito *et al.*, 2000; Locke, 1998; Locke *et al.*, 2000; Nicolson *et al.*, 1988; Saunders *et al.*, 2001; el-Sabban y Pauli, 1991; el-Sabban y Pauli, 1994-1995). Se ha demostrado que una posible función supresora de BRMS1 sobre metástasis es mediada a través de la regulación de la expresión de genes codificantes para conexinas, ejerciendo un efecto de restauración sobre las comunicaciones intercelulares (Saunders *et al.*, 2001).
- **Miembro del complejo mSin3/HDAC**. El complejo multiproteico co-represor de la transcripción formado por mSin3A (“Suppressor of defective silencing 3”) y por Histonas Desacetilasas (mSin3/HDAC) desarrolla un papel fundamental en el modelaje de la cromatina y con ello en la regulación de la transcripción. Existen diversos artículos que describen la interacción de BRMS1 con diferentes miembros de dicho complejo (Doyon *et al.*, 2006; Hurst *et al.*, 2008; Liu *et al.*, 2006; Meehan *et al.*, 2004; Nikolaev *et al.*, 2004). Recientemente se ha descrito un nuevo miembro, la proteína p40, que mantiene elevada homología con BRMS1. Tanto p40 como BRMS1 son capaces de inhibir crecimiento celular mediante represión de la transcripción de genes asociados al crecimiento celular, posiblemente al reclutar el complejo de histonas desacetilasas (HDAC) (Nikolaev *et al.*, 2004).
- **Inhibición de la ruta de NFκB**. Una de las rutas por las que BRMS1 regula indirectamente la transcripción de genes puede ser la del factor de transcripción NFκB (Cicek *et al.*, 2005; Liu *et al.*, 2006; Yang *et al.*, 2008) a través de diferentes vías aún desconocidas. Además, se ha descrito un efecto directo de BRMS1 sobre el estado de activación de una de las subunidades del factor NFκB, a nivel de

modificaciones post-traduccionales como es la acetilación/desacetilación, pudiendo modificar por lo tanto su capacidad transcripcional (Chen *et al.*, 2002; Kiernan *et al.*, 2003; Liu *et al.*, 2006; Samant *et al.*, 2007).

- ***Amplio abanico de interacciones proteicas.*** Además de interaccionar con las histonas desacetilasas tipo I y II (HDAC1 y HDAC2) y formar parte del complejo en el que actúan (mSin3/HDAC), recientemente se ha descrito la interacción de la proteína BRMS1 con HDAC3 (Samant *et al.*, 2007), que controla la expresión de osteopontina (OPN) a través del sitio de unión de NFκB al promotor de OPN. HDAC3 desacetila la subunidad p65 dando lugar a una disminución en la actividad de NFκB (Chen *et al.*, 2001), por lo que hay autores que sugieren que es BRMS1 quién desacetila p65 de NFκB posiblemente por su asociación con HDAC3, traduciéndose en una supresión de la expresión de OPN (Samant *et al.*, 2007).

De igual manera, se ha descrito la interacción con otras proteínas, tales como NMI (“N-myc interactor”), MRJ (Chaperona relacionada con Hsp40), CCG1 (o TAFII250, proteína esencial para la progresión de la fase G), SMTN (“Smoothelin”, proteína específica del citoesqueleto en músculo liso), BAF 57 (Factor asociado a BRG1) y la proteína de estrés térmico Hsp90, que se ha descrito estabiliza a BRMS1 (Hurst *et al.*, 2006; Hurst *et al.*, 2008), si bien la mayoría de estas interacciones no están comprobadas.

Otra proteína cuya interacción con BRMS1 se ha demostrado, es la proteína 1 de unión a retinoblastoma (RBP1 ó ARID4A) (Meehan *et al.*, 2004), que al unir al supresor de tumor retinoblastoma (RB) (Fattaey *et al.*, 1993), lo recluta hacia promotores dependientes del factor de transcripción E2F (Lai *et al.*, 1999a; Lai *et al.*, 1999b), inhibiendo la expresión de genes dependiente de E2F, y suprimiendo crecimiento celular.

A pesar de ser una proteína supresora de metástasis identificada como tal hace poco tiempo, muchos son los trabajos científicos que se están desarrollando y publicando en la actualidad que relacionan BRMS1 con cambios en los patrones de expresión de diversas proteínas de mucha relevancia biológica y con funciones ya conocidas. Se han realizado estudios de proteómica que muestran grandes cambios en la expresión de genes cuando varían los niveles de BRMS1. Un ejemplo de ello son los experimentos realizados recientemente por el Dr. Rivera, en los que tras sobreexpresar y silenciar BRMS1 en líneas celulares de melanoma, observaron patrones distintos de expresión diferencial. Tras realizar análisis mediante espectrometría de masas, identificaron 43 proteínas diferencialmente expresadas cuando BRMS1 era sobreexpresado, de las que más del 75% estaban inhibidas. Los niveles de todas las proteínas identificadas en células en las que se había anulado la expresión de BRMS1 eran mas elevados que los de las células sin tratar, además de acontecerse cambios fenotípicos en la célula, una considerable disminución en la formación de lamelipodios y una morfología mas elongada. Muchas de las proteínas diferencialmente expresadas están involucradas en procesos de gran importancia biológica como crecimiento celular, procesos enzimáticos, rutas metabólicas y de señalización (Rivera *et al.*, 2007).

Por otro lado, los dos mecanismos que cobran mayor importancia en los últimos estadios de la cascada metastásica, son la supervivencia celular a través del torrente

circulatorio y linfático, y cambios en la adhesión que facilitan la migración celular a la vez que esta supervivencia. Por ello, es importante estudiar la relación que BRMS1 pueda tener en el desarrollo de estos procesos.

La “apoptosis” o “muerte celular programada” es el principal mecanismo para combatir la supervivencia celular acontecida en el desarrollo metastásico. Existen indicios sobre la participación de la proteína BRMS1 en la regulación de rutas que regulan expresión de genes relacionados con apoptosis, tanto inhibiendo la expresión de genes antiapoptóticos, regulados a través de la vía clásica de NFκB (cIAP-2, Bfl-1/A1 y Bcl-xL) (Liu *et al.*, 2006), como regulando la de genes pro-apoptóticos (Bim y Bmf) (Phadke *et al.*, 2008). De este modo, BRMS1 podría sensibilizar a las células cancerígenas hacia muerte celular programada o apoptosis.

Quizás en lo que se refiere a la supervivencia o no de células en suspensión que mediante el torrente circulatorio se dirigen a focos distantes, sería importante referirnos a “Anoikis”, o muerte celular programada independiente de sustrato (Frisch y Francis, 1994). Usando un modelo de células tumorales (NSCLC) en suspensión, se demostró que aquellas expresando transitoriamente BRMS1 tenían significativamente índices de muerte mas elevados que el control. Por otro lado, células en las que se había silenciado la expresión de BRMS1, sobrevivían en mayor número y mas tiempo en suspensión (Liu *et al.*, 2006). Por lo tanto, podemos concluir que BRMS1 inhibe la supervivencia, al igual que el crecimiento y la proliferación de este foco secundario (Hedley *et al.*, 2008; Nikolaev *et al.*, 2004).

En lo que respecta a efectos de BRMS1 sobre adhesión celular, existen diferentes puntos de vista. Unos autores explican que no hay evidencias que muestren que BRMS1 regule la progresión metastásica a través de alteraciones en la adhesión, bien sea por cambios en componentes de la matriz extracelular, o en expresión de metaloproteasas de matriz o heparanasas, si bien sí demuestran que BRMS1 inhibe migración inducida por la quimioquina SDF-1α de células tumorales al reducir la expresión del Receptor 4 de quimioquina CXCR4 por la vía de NFκB (Yang *et al.*, 2008). Sin embargo, otros autores indican mediante estudios “*in vitro*”, que BRMS1 inhibe claramente la adhesión de células al sitio secundario, además de inhibir supervivencia bajo estrés (hipoxia) y crecimiento de colonias en un ambiente independiente de anclaje (Hedley *et al.*, 2008).

Con estos ensayos se ha podido concluir que al menos dos etapas en la cascada metastásica están inhibidas por BRMS1, tanto la supervivencia celular inicial como la iniciación del crecimiento de micrometástasis, etapas en la cascada que clínicamente representan prósperas dianas terapéuticas, ampliando el abanico de posibilidades en investigación.

3. *Tax1BP1*

Esta proteína fue identificada en 1995 como proteína de unión a la oncoproteína Tax, del virus humano HTLV-1 (“Human T-cell leukemia virus type-I”) (Gachon *et al.*, 1998), que fue el primer retrovirus asociado a una enfermedad humana, la leucemia adulta de las células T (Matsuoka, 2005), y también está relacionada con la enfermedad neuroinflamatoria HAM-TSP.

La proteína Tax es un factor de transcripción que tiene como diana diferentes promotores virales heterólogos y una variedad de genes celulares (Yoshida, 2001). Tax también actúa a través de elementos de unión a ADN que son reconocidos por el factor de transcripción celular CREB (“cAMP Responsive Element Binding protein”, de respuesta a cAMP).

Se ha descrito que Tax activa el factor de transcripción NF κ B (Yoshida, 2001), cuyos niveles se presentan también elevados en ratones defectivos de Tax1BP1 (Iha *et al.*, 2008). Por otro lado, se ha descrito que Tax también tiene actividad represora sobre la expresión de determinados genes celulares, entre los que se encuentran la polimerasa β de ADN (Jeang *et al.*, 1990), p53 (Pise-Masison *et al.*, 1998; Van *et al.*, 2001), ciclina A (Kibler y Jeang, 2001), c-Myb (Nicot *et al.*, 2001), MyoD (Riou *et al.*, 2000), factor de crecimiento transformante β (Arnulf *et al.*, 2002) y receptores nucleares (Doucas y Evans, 1999).

Tax1BP1, sin embargo, es un coactivador transcripcional de receptores nucleares, y recientemente se ha demostrado que Tax inhibe la función de receptores nucleares a través de su unión a Tax1BP1, induciendo su disociación del receptor (Chin *et al.*, 2007).

La pérdida de Tax1BP1 (en ratones defectivos de esta proteína) se ha relacionado directamente con el desarrollo de una enfermedad llamada valvulitis inflamatoria cardíaca, que desemboca en muerte prematura y desarrollo de hipersensibilidad en animales tratados con bajas dosis de TNF- α e IL-1 β , estando sobre-activada la ruta de NF κ B (Iha *et al.*, 2008).

3.1 *Características estructurales de Tax1BP1*

Tax1BP1 es una proteína cuyo patrón de expresión es muy ubicuo y se localiza principalmente en el núcleo.

Existen tres isoformas, originadas por procesamiento alternativo, siendo la isoforma 1 la mas representada y de la que mas información existe. Esta isoforma (Q86VP1-1) comprende 789 aminoácidos y tiene un peso molecular de 90.877 Da. Además se han descrito tres variantes diferentes de la isoforma 1 de Tax1BP1 que se basan en mutaciones aminoacídicas puntuales, que afecta a las regiones centrales (Ling y Goeddel, 2000).

En su región amino terminal (aminoácidos 1 a 150) existe una secuencia homóloga a la proteína nuclear NDP52, seguido de tres motivos “coiled-coil”, en los segmentos 185-248, 350-413 y 526-589. Además, contiene un dominio predicho típico de oligomerización entre los residuos 320 y 340. En el extremo carboxilo terminal se

localizan dos dominios potenciales de unión a ubiquitina (*UBZs*), según estudios publicados recientemente (Iha *et al.*, 2008), que están estructurados en “dedo de zinc” del tipo Cys₂/His₂ (730-753 y 757-780) (De Valck *et al.*, 1999; Ling y Goeddel, 2000). El segundo de los dominios (*UBZ2*) pertenece a la familia de motivos en “dedos de zinc” de unión a ubiquitina de polimerasa η (Bienko *et al.*, 2005) (Figura 4).

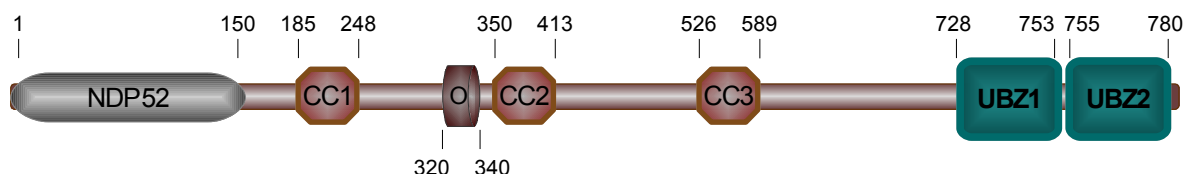


Figura 4. Representación esquemática de la predicción de motivos de la proteína Tax1BP1 humana. Se muestran en números los aminoácidos delimitantes de todos los motivos estructurales. Desde el extremo amino al carboxilo terminal, Tax1BP1 contiene: una región homóloga a NDP52, tres motivos “coiled-coil” (nombrados como CC1, CC2 y CC3), un motivo con homología con dominios de oligomerización (“O”) y dos dominios potenciales de unión a Ubiquitina (UBZ1 y UBZ2) en estructura de “dedos de zinc”, del tipo Cys₂/His₂. Figura realizada a partir de información extraída de diferentes bases de datos y herramientas de predicción informáticas (Referencias 1-11).

Ambos dominios contienen secuencias consenso “PPxY”, donde “x” es una asparagina (N) y un aspártico (D) en *UBZ1* y *UBZ2* respectivamente, responsables de la interacción con proteínas que contienen dominios con dos triptófanos conservados (*WW*) (Sudol *et al.*, 1995). Los dos dominios *UBZ* presentan un elevado grado de conservación entrediferentes especies, como se puede observar en la Figura 5.

	728	PPxY ₁	753	755	PPxY ₂	780
<i>Homo sapiens</i>	KKCP	LC	ELMF	PPNYDQSKFEEHVESH	WKVCPMCSEQFP	PPDYDQQVFERHVVQTH
<i>Canis lupus familiaris</i>	KKCP	LC	ELMF	PPNYDQSKFEEHVESH	WKVCPMCSEQFP	PPDYDQQGFERHVVQTH
<i>Pan troglodytes</i>	KKCP	LC	ELMF	PPNYDQSKFEEHVESH	WKVCPMCSEQFP	PPDYDQQVFERHVVQTH
<i>Mus musculus</i>	KKCP	LC	ELMF	PPNYDQTKFEEHVESH	WKVCPMCSEQFP	PPDYDQQGFERHVVQTH
<i>Gallus gallus</i>	KKCP	LC	DLMF	PPNYDQSKFEEHVESH	WKVCPMCSEQFP	PPDYDQQGFERHVVQTH
	Dedo Zn 1				Dedo Zn 2	

Figura 5. Alineamiento de secuencia de los dominios estructurados en “dedos de zinc” de los homólogos de Tax1BP1 humano en diferentes especies. Con números se señalan los aminoácidos delimitantes de estos dominios. Las secuencias consenso PPxY entre los dos *UBZ* se muestran recuadrados, siendo x=N y x=D en los dominios de unión a zinc 1 y 2 respectivamente.

La secuencia aminoacídica completa de Tax1BP1 (a excepción de los dos “dedos de zinc”) mantiene un 27% de identidad y 45% de similitud con una proteína recientemente identificada, un coactivador transcripcional rico en motivos “coiled-coil”, CoCoA (Kim *et al.*, 2003; Yang *et al.*, 2006), que forma parte del complejo co-activador del receptor nuclear p160.

3.2 Ligandos celulares de Tax1BP1 y sus implicaciones en procesos de ubiquitinación

Se han descrito diversas interacciones entre Tax1BP1 y distintos elementos importantes en vías de regulación de transcripción génica, en rutas de señalización de respuestas inflamatorias, de muerte celular, etc., y aunque los mecanismos que acontecen para llevar a cabo su acción están aún por descubrir, a continuación se enumeran sus interacciones con mayor relevancia biológica.

3.2.1 Tax1

Los primeros indicios de la interacción entre Tax1BP1 y Tax fueron obtenidos mediante ensayos de doble híbrido en levadura (Gachon *et al.*, 1998; Wu *et al.*, 2004). Posteriormente esta interacción fue comprobada mediante ensayos de coimmunoprecipitación y de obtención de complejos por precipitación usando resina de GST (Chin *et al.*, 2007).

Hace algunos años, Doucas y Evans demostraron la capacidad de la proteína Tax de inhibir la actividad transcripcional de varios receptores nucleares, entre los que se incluirían el receptor de glucocorticoide (Doucas y Evans, 1999). Estos receptores nucleares son factores de transcripción que regulan la expresión de genes en respuesta a esteroides, retinoides y otras moléculas de señalización, y juegan un papel fundamental en el desarrollo, diferenciación, metabolismo y proliferación celular, de forma que la desregulación de la transcripción dependiente de estos receptores conduce al desarrollo de enfermedades, tales como cáncer (Fu *et al.*, 2004). En base a este estudio, se demostró que Tax1BP1 es un coactivador transcripcional de receptores nucleares, y que Tax reprime la función de estos receptores a través de la unión con Tax1BP1, disociándolo del receptor (Chin *et al.*, 2007).

Se ha descrito la interacción de Tax con la subunidad reguladora IKK γ (NEMO) de NF κ B (Jin *et al.*, 1999), lo que representaría la evidencia mas directa del papel de Tax sobre dicha ruta, lo que le conduce a ser catalogado como un potente activador de NF κ B (Ballard *et al.*, 1988; Leung y Nabel, 1988). Para ello interacciona con NEMO y promueve la activación de IKK, que a su vez fosforila a I κ B α , permitiendo la activación del factor de transcripción (Chu *et al.*, 1999; Harhaj y Sun, 1999; Jin *et al.*, 1999). Aunque no se conoce muy bien la forma de actuación de Tax en la ruta de NF κ B, se cree que Tax1BP1 captura a Tax inhibiendo su función de activación sobre NF κ B, por lo que la formación del complejo reprimiría la ruta.

3.2.2 A20

A20 es una proteína que inhibe la muerte celular programada mediada por el factor de necrosis tumoral. Ha sido caracterizada como un regulador negativo de procesos de inflamación (Krikos *et al.*, 1992; Lee *et al.*, 2000), e inhibidor de la activación de NF κ B por mecanismos aún desconocidos (Jäättelä *et al.*, 1996; Wertz *et al.*, 2004). Es una

proteína de 790 residuos que contiene siete dedos de zinc del tipo Cys₂/Cys₂, y es inducida por una variedad de estímulos inflamatorios.

La interacción entre Tax1BP1 y A20 ha sido rescatada mediante ensayo de doble híbrido en levadura, determinando que esta interacción está mediada por la mitad carboxilo terminal de la proteína A20 (residuos 373-790), que incluye los siete “dedos de zinc” (De Valck *et al.*, 1999), no estando involucrados en tal interacción los dominios “dedos de zinc” de Tax1BP1 (Shembade *et al.*, 2008). Sin embargo, existe una controversia con los resultados obtenidos por otros autores, que concluyen que el segundo UBZ de Tax1BP1 media su unión con A20 (Iha *et al.*, 2008).

El mecanismo de apoptosis inducido por TNF o CD95 se asocia con la proteólisis de Tax1BP1 mediada por caspasas. En casos similares donde existe degradación por parte de caspasas se generan fragmentos que llevan a la célula a muerte programada, como es el caso de las proteínas antiapoptóticas Bcl-2 y Bcl-X_L (Fujita *et al.*, 1998). Se ha sugerido que Tax1BP1 media el efecto antiapoptótico de A20 y que es procesado por caspasas durante la apoptosis (De Valck *et al.*, 1999).

Como mencionamos anteriormente, se ha descrito la capacidad de la proteína A20 de inhibir la expresión de genes dependiente de NFκB inducida por TNF, relacionándose con esta ruta gracias a su actividad de desubiquitinación, directamente dirigida hacia las lisinas 48 y 63 unidas en las cadenas de poliubiquitina. Además, la fosforilación de A20 por parte de IKKβ potencia su capacidad de inhibir la ruta de NFκB (Hutti *et al.*, 2007). Como Tax1BP1 interacciona con A20, la inhibición por parte de Tax1BP1 sobre la ruta de NFκB es a través de la regulación de la función de A20 (Shembade *et al.*, 2007), englobando procesos de ubiquitinación/desubiquitinación detallados mas adelante.

3.2.3 TRAF6

La asociación entre Tax1BP1 y TRAF6 (“TNF receptor associated factor 6”) fue descrita mediante ensayos de coimmunoprecipitación (Ling y Goeddel, 2000). TRAF6 es una proteína adaptadora, que interviene en la ruta de señalización de NFκB inducida por IL-1β, y es esencial para la activación del factor de transcripción (Cao *et al.*, 1996).

Se ha demostrado que la formación del complejo Tax1BP1-TRAF6 depende de la presencia de la quinasa asociada al receptor de interleuquina 1 (IRAK, “IL-1 receptor-associated kinase”), y que la interacción entre estas dos proteínas ocurre a través del motivo “coiled-coil” de Tax1BP1 y el motivo en anillo (en inglés “RING finger motif”) de la región amino terminal sumado a los dominios “dedos de zinc” de TRAF6 (Ling y Goeddel, 2000). Estos resultados contrastan con otro estudio en el cual sus autores avalan que los dominios “dedos de zinc” de Tax1BP1 son imprescindibles para su interacción con TRAF6, ya que la mutación de los mismos, eliminaba la asociación entre las dos proteínas (Iha *et al.*, 2008).

Se ha descrito la existencia de un estado ubiquitinado de TRAF6 en la activación de la ruta de NFκB. Por otro lado, se ha observado una correlación entre la sobreexpresión de Tax1BP1 y una fuerte reducción de la poliubiquitinación de TRAF6, que estaría mediada por la interacción de Tax1BP1 con la proteína A20 (Iha *et al.*, 2008). De esta forma se inhibiría la acción positiva de TRAF6 sobre la vía.

3.2.4 *RIP1*

RIP1 (“Receptor Interacting Protein”) es una proteína serín-treonín quinasa. Esta proteína tiene capacidad de unir ubiquitinas, estado que se ha descrito está relacionado con la activación de la ruta de NFκB. Es importante en la transducción de señales, siendo un regulador positivo de la cascada de señalización de NFκB.

La interacción entre Tax1BP1 y RIP1 ha sido descrita recientemente mediante ensayos de coimmunoprecipitación, demostrando que los dominios estructurados en “dedos de zinc” de Tax1BP1 son imprescindibles para su unión con RIP1 (Shembade *et al.*, 2008), y además la sobreexpresión de Tax1BP1 inhibe la poliubiquitinación de RIP1, inhibiendo con ello la ruta de señalización de NFκB (Iha *et al.*, 2008).

3.2.5 *Itch (ligasa E3)*

Itch o ligasa E3 de ubiquitina es una proteína que contiene cuatro dominios ricos en pares de triptófanos (*WW*; triptófano-triptófano). Ensayos de coimmunoprecipitación realizados con diversos mutantes de la proteína Tax1BP1 han mostrado que su interacción con la proteína Itch está mediada por los dominios *PPxY* de Tax1BP1 (Shembade *et al.*, 2008), apoyando la premisa de que los dominios *PPxY* unen dominios *WW* (Sudol *et al.*, 1995).

Analizando en conjunto todas las interacciones, observamos que Tax1BP1 forma parte de un complejo multiproteico, integrado por las proteínas Itch, Tax1BP1, y A20, que parece inhibir la ubiquitinación de TRAF6 y RIP1. La des-ubiquitinación de ambas proteínas tiene un efecto negativo sobre procesos inflamatorios y de activación de la transcripción mediada por NFκB. Además, la formación de dicho complejo multiproteico está inhibida por Tax (Shembade *et al.*, 2008), que representa otra relación entre Tax1BP1 y Tax.

Los procesos de ubiquitinación en rutas de señalización celulares son muy importantes. Tanto moléculas de ubiquitina, como proteínas tipo ubiquitinas, actúan como mensajeros de señales que controlan muchas funciones celulares, como proliferación celular, apoptosis, ciclo celular y reparación del ADN. Así, las alteraciones en las vías de ubiquitinación resultan en el desarrollo de enfermedades humanas, incluyendo el desarrollo de tumores (Haglund y Dikic, 2005; Hoeller *et al.*, 2006).

4. *BRMS1 y Tax1BP1, reguladores negativos de la activación de NFκB*

Muchas son las vías de señalización involucradas en el desarrollo de metástasis. Una de ellas es la activación de la ruta del factor de transcripción NFκB, en la que existe un elevado número tanto de inductores como de inhibidores, que resultan en la inhibición y activación de apoptosis respectivamente. Las dos proteínas expuestas anteriormente, *BRMS1*, un supresor de metástasis, y *Tax1BP1*, proteína de unión a la oncoproteína Tax1,

son efectores negativos de la ruta de activación de NFκB (Cicek *et al.*, 2005; Shembade *et al.*, 2007). En este apartado, intentaremos acercarnos a los mecanismos que relacionan estas dos proteínas con dicha ruta.

4.1 Aspectos moleculares de la ruta de activación de NFκB

NFκB (en inglés “Nuclear Factor-kappa B”) es un factor de transcripción heterodimérico involucrado en respuestas inflamatorias (Barnes y Karin, 1997), y a estímulos como estrés y agentes citotóxicos (Gilmore, 1999), regulando la expresión de genes implicados en invasión, metástasis, angiogénesis (Karashima *et al.*, 2003) y apoptosis celular (Barkett y Gilmore, 1999; Gapuzan *et al.*, 2002; Karin, 2006; Lin y Karin, 2003; Sliva *et al.*, 2002). Por ello, su activación constitutiva está relacionada con diferentes tipos de cáncer (Karin, 2006; Lin y Karin, 2003).

Existen tres vías fundamentales para la activación de NFκB (Figura 6). Las dos primeras que se describen a continuación, se conocen con mas profundidad y están mejor caracterizadas (Gilmore, 2006).

1. La “*vía clásica*” o “*canónica*”, en la que el dímero formado por p50 y RelA se mantienen en el citoplasma en estado inactivo debido a su interacción con una molécula inhibidora IκB. Normalmente este inhibidor es IκBα.

Tras la unión de un ligando a un receptor en la superficie celular, tales como el de factor de necrosis tumoral o un receptor tipo Toll, se reclutan adaptadores como TRAF y RIP en el dominio citoplasmático del mismo. Estos adaptadores a su vez, reclutan el complejo IKK, que contiene las subunidades catalíticas α y β, así como dos moléculas reguladoras γ o NEMO (“NFκB Essential Modulator”). Esta interacción suele estar mediada por la capacidad de NEMO de unir cadenas de ubiquitinas enlazadas por sus lisinas 63.

La formación de este complejo macromolecular provoca la activación de IKK, que fosforila a IκB en dos residuos de serina y provoca su ubiquitinación (de ubiquitinas unidas por sus lisinas 48) y posterior degradación proteosomal. Esta acción conduce a la translocación del dímero p50/RelA al núcleo, donde reconoce secuencias específicas de ADN, activando la transcripción de determinados genes.

2. La “*vía alternativa*” o “*no canónica*”, está mediada por la activación del complejo p100/RelB durante el desarrollo de células B y T. Esta ruta difiere de la clásica en que solamente ciertas señales de receptor, como por ejemplo LTb (“Lymphotoxin B”), BAFF (“B-cell activating factor”) y CD40, son capaces de activar la vía. En dicha ruta de activación, el complejo IKK solamente está formado por sus dos subunidades catalíticas α, mientras que NEMO está ausente.

La unión de un adaptador al receptor conduce a la activación de una quinasa (NIK, en inglés “NFκB-inducing kinase”), que fosforila y activa el complejo IKKα, provocando la fosforilación de dos residuos de lisina localizados en el extremo carboxilo terminal del dominio IκB de p100, conduciendo a su proteólisis parcial y liberación del complejo p52/RelB.

3. Se ha postulado la existencia de una tercera ruta, donde no se conoce aún la forma de regulación. En ella, homodímeros de p50 (o de p52) se translocarían al núcleo y llegarían a ser activadores de la transcripción al unirse a factores como Bcl-3 (“B cell leukemia/lymphoma 3”) (Franzoso *et al.*, 1992).

En estas tres rutas de activación del factor de transcripción NFκB, tienen lugar diferentes modificaciones postraduccionales de las subunidades de NFκB como son fosforilación, acetilación y proil-isomerización, que pueden modular la actividad transcripcional del factor de transcripción, contribuyendo en cada etapa de la vía (Perkins, 2006). Por ejemplo, la ubiquitinación o fosforilación de la subunidad IκBα en determinados residuos conducen a su degradación (Hayden y Ghosh, 2004), o la acetilación de RelA potencia su unión al ADN e inhibe su asociación con IκBα (Chen *et al.*, 2002). También se da una estrecha relación entre las diferentes modificaciones postraduccionales, tal es el caso de fosforilación y acetilación, cuando la fosforilación en dos serinas de RelA potencia su unión a la subunidad p300, lo que lleva a acetilar RelA en otros residuos y lo activa (Chen *et al.*, 2005; Hoberg *et al.*, 2004).

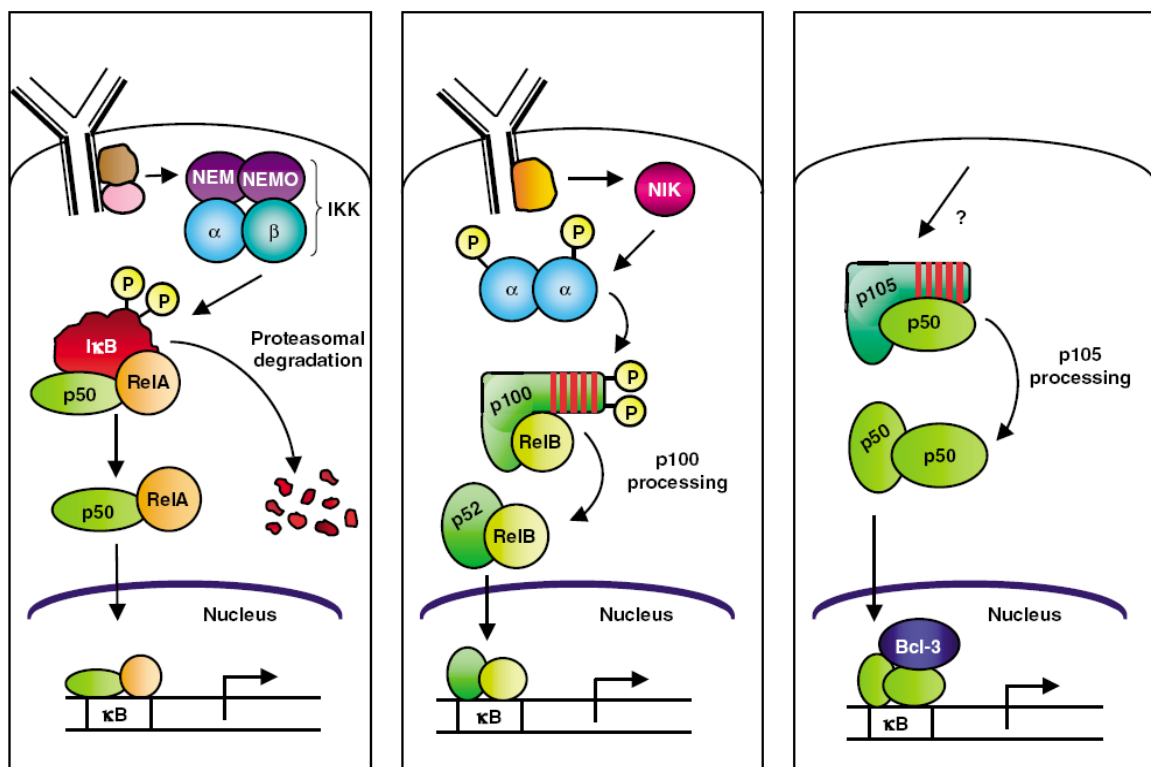


Figura tomada de TD Gilmore, *Oncogene* (2006) (Gilmore, 2006)

Figura 6. Rutas de señalización del factor de transcripción NFκB (Gilmore, 2006). **a)** La “vía clásica” o “canónica”, en la que el dímero formado por p50/RelA se mantiene en el citoplasma en estado inactivo unido a una molécula inhibidora IκB. Bajo la señalización desencadenada por la unión de ligando-receptor se recluta el complejo IKK (con dos subunidades catalíticas α, β y reguladoras γ o NEMO) que va a fosforilar IκB en dos residuos de serina, y conduce a su ubiquitinación y consecuente degradación proteasomal, dejando NFκB en su forma activa, que se transloca al núcleo para activar la transcripción génica. **b)** La “vía alternativa” o “no canónica”, en la que la unión del ligando al receptor conduce a la activación de una quinasa inductora de NFκB (NIK) que fosforila y activa el complejo IKKα, que a su vez fosforila el dominio IκB de p100, conduciendo a su proteólisis parcial y liberación del complejo p52/RelB

activo que se transloca al núcleo para activar la transcripción de genes. c) Una tercera ruta, donde no se conoce aún la forma de regulación. Homodímeros de p50 (o de p52) se translocan al núcleo, y llegan a ser activadores de la transcripción al unirse a factores nucleares como Bcl-3.

4.2 BRMS1 como modulador de la ruta de activación de NFκB

Se ha descrito el efecto negativo que BRMS1 ejerce sobre la expresión de diferentes genes. A continuación se relacionan algunos genes, que además de estar controlados a través de la vía de NFκB, su expresión está negativamente regulada por la acción de BRMS1 y están implicados en mecanismos desregulados en el proceso de metástasis.

- ✓ **Regulación negativa sobre el Activador de plasminógeno del tipo Uroquinasa (uPA).** Un factor importante a tener en cuenta en la etapa de diseminación metastásica es la proteólisis y degradación de la matriz extracelular, así como la migración celular. Uno de los mediadores del remodelado de esta matriz extracelular es el activador de plasminógeno uPA, una serín-proteasa que estimula la conversión de plasminógeno inactivo a plasmina (Liotta *et al.*, 1981; Quigley, 1979). Se ha demostrado que esta proteasa juega un papel importante en el desarrollo de metástasis (Kim *et al.*, 1998), al igual que se ha demostrado que elevados niveles de uPA y de su inhibidor PAI-1 (“Plasminogen activator inhibitor 1”), son indicadores de mal pronóstico en pacientes con cáncer de mama (Grøndahl-Hansen *et al.*, 1993; Jänicke *et al.*, 1989; Look *et al.*, 2003; Zemzoum *et al.*, 2003).

Estudios recientes han demostrado que NFκB activa la expresión de uPA, así como de otros genes implicados en angiogénesis y metástasis (Barnes y Karin, 1997; Karashima *et al.*, 2003; Sliva *et al.*, 2002). Se ha podido verificar una correlación entre sobre-expresión de BRMS1 y reducida expresión de la proteína uPA. Además, se ha comprobado un efecto inhibitorio de BRMS1 sobre la actividad de uPA inducida por TNF-α (Cicek *et al.*, 2005). Aunque el factor de transcripción AP-1 (“Adaptor Protein 1”) es otro elemento que también regula la expresión del gen uPA (Dunn *et al.*, 2001), se ha demostrado que la expresión de BRMS1 no ejerce ningún efecto sobre la capacidad de unión de AP-1 al ADN sugiriendo que BRMS1 específicamente suprime la expresión de uPA a través de la vía de NFκB (Cicek *et al.*, 2005).

- ✓ **Regulación negativa sobre la Osteopontina (OPN).** La osteopontina es una fosfoglicoproteína de secreción extracelular que se asocia con progresión tumoral. La línea celular de mama MDA-MB-435 es un ejemplo de células que secretan elevados niveles de OPN. En el año 2007, se mostró cómo la expresión de OPN disminuía hasta un 95% en células MDA-MB-435 transfectadas con BRMS1 (Samant *et al.*, 2007). Además, consiguieron identificar un nuevo sitio de unión del factor de transcripción NFκB al promotor de OPN, en cuya regulación se ha demostrado estar implicada la histona desacetilasa 3 (HDAC3). Así, sugirieron que la represión que BRMS1 ejerce sobre la expresión de OPN es desacetilando la

subunidad p65 de NFκB, posiblemente por su asociación con HDAC3 (Samant *et al.*, 2007).

- ✓ *Inhibición de la expresión del Receptor 4 de quimioquina C-X-C (CXCR4).* En las etapas de intravasación y extravasación que se suceden en el proceso de desarrollo metastásico es donde la quimiotaxis juega un papel importante, y con ello la vía de señalización de SDF-1α/CXCR4 (“Stromal cell Derived Factor 1α/Chemokine, CXC motif, receptor 4”). Mediante estudios celulares y de quimiotaxis en una línea celular de pulmón (NSCLC), concluyeron que BRMS1 inhibe migración de células tumorales inducida por SDF-1α, al reprimir la expresión de CXCR4 dependiente de NFκB (Yang *et al.*, 2008).
- ✓ Con el fin de determinar si BRMS1 inhibe la expresión de determinados genes antiapoptóticos regulados a través de la vía clásica de NFκB, se ha estudiado el efecto de sobreexpresar BRMS1 sobre los niveles de estos genes, observando que BRMS1 significativamente reduce la expresión de los genes antiapoptóticos cIAP-2, Bfl-1/A1 y Bcl-xL, dependientes de esta ruta (Liu *et al.*, 2006).

4.2.1 *Modificaciones postraduccionales sobre el factor de transcripción NFκB por parte de BRMS1*

Se ha demostrado que existe interacción directa entre BRMS1 y la subunidad RelA (p65) del factor de transcripción NFκB. Concretamente, la unión se da entre BRMS1 y la región comprendida entre los residuos 245 y 354 de la subunidad, que no incluye ninguno de los tres dominios de transactivación implicados en su unión con diferentes componentes transcripcionales basales y coactivadores transcripcionales (Liu *et al.*, 2006).

Se ha visto una directa regulación negativa por parte de BRMS1 sobre varias etapas de la regulación en la activación de NFκB:

✓ *Fosforilación de IκBα:*

Diversos autores han demostrado que BRMS1 inhibe la fosforilación y consecuente degradación de IκBα, impidiendo la activación del factor de transcripción NFκB y por tanto su translocación al núcleo para allí activar la transcripción de determinados genes (Cicek *et al.*, 2005; Liu *et al.*, 2006) (Figura 7).

✓ *Acetilación de NFκB:*

Otras modificaciones post-traduccionales sobre el factor de transcripción NFκB, como es la acetilación/desacetilación, son importantes para activar su capacidad de unión al ADN y con ellos su actividad transcripcional, de forma que se ha observado que si NFκB se encuentra acetilado se potencia su activación (Chen *et al.*, 2002; Kiernan *et al.*, 2003).

Se ha comprobado una relación inversa entre niveles de la proteína BRMS1 y el estado de acetilación de la subunidad RelA/p65 de NFκB. Algunos experimentos realizados recientemente han verificado que BRMS1 inhibe la acetilación de la Lisina 310 de la subunidad RelA/p65. Tras bloquear la expresión de BRMS1 mediante el uso de

ARNsi (ARN de interferencia), se produce un aumento drástico de la acetilación endógena de RelA/p65, además de una dramática disminución en la unión de HDAC1 con RelA/p65, requiriéndose por lo tanto la proteína BRMS1 para la desacetilación del factor de transcripción. Se han realizado ensayos de acetilación/desacetilación *in vitro* para estudiar si esta capacidad de desacetilación sobre RelA/p65 es proporcionada por una posible actividad desacetilasa intrínseca de BRMS1 o bien si es debido a histonas desacetilasas. Finalmente se concluyó que posiblemente ambas proteínas BRMS1 y HDAC1 funcionan conjuntamente para regular la desacetilación de RelA/p65 (Liu *et al.*, 2006) (Figura 7).

4.3 Relevancia de Tax1BP1 en la ruta de activación de NFκB

Tax1BP1, se caracteriza por ser un coactivador de receptor nuclear, a su vez regulador negativo de la activación de NFκB inducida por TNF-α y por IL-1-β.

La participación de Tax1BP1 en el complejo multiproteico involucrado en mecanismos de ubiquitinación, sugiere que éste es el principal mecanismo por el cual Tax1BP1 inhibe la ruta de NFκB (Iha *et al.*, 2008; Shembade *et al.*, 2007; Shembade *et al.*, 2008).

4.3.1 Procesos de ubiquitinación en regulación de la ruta NFκB

Los procesos de ubiquitinación en rutas de señalización, como hemos mencionado anteriormente, son muy importantes, comprendiendo señales tanto de activación como de represión sobre diferentes moléculas que hacen que la ruta celular se dirija hacia una dirección u otra (Wullaert *et al.*, 2006). Recientemente se han desarrollado estudios que caracterizan el proceso de ubiquitinación como un mecanismo para la activación de NFκB (Chen, 2005; Haglund y Dikic, 2005), y lo que es clínicamente mas importante, como un mecanismo relacionado con cáncer (Hoeller *et al.*, 2006). Existen estudios que demuestran que la ubiquitinación regula al menos tres etapas en la vía de activación de NFκB: en la degradación del inhibidor de la ruta, IκB, en el procesamiento de precursores de NFκB y en la activación del activador de IκB, IKK (Chen, 2005). Ya que existe una estrecha relación entre la desregulación de la ruta de activación de NFκB y la progresión a metástasis, una función evidente que relaciona Tax1BP1 con metástasis, es a través de los mecanismos de ubiquitinación sobre diferentes moléculas involucradas también en metástasis.

Además de estar descritas las interacciones anteriormente citadas con Tax1BP1 (en el apartado “3.2”), también se han demostrado las establecidas entre las demás proteínas entre sí; esto es, entre A20-TRAF6 (Heyninck y Beyaert, 1999), A20-Tax, A20-RIP1 (Wertz *et al.*, 2004), Itch-A20, Itch-RIP1 (Shembade *et al.*, 2008), Itch-TRAF6 (Shembade *et al.*, 2008), etc. Esto hace entrever la existencia de un complejo multiproteico involucrado en procesos de ubiquitinación, regulado por la presencia de A20, y que actualmente está siendo estudiado por diversos grupos (Figura 7). Además, la formación de dicho complejo está reprimida por la oncoproteína Tax, que inhibe las interacciones entre Tax1BP1, Itch y A20 (Shembade *et al.*, 2008).

Se ha demostrado que la molécula reguladora Tax1BP1 actúa como una proteína adaptadora que recluta la ligasa de ubiquitina E3 hacia la proteína A20 mediante la interacción de motivos *PPxY-WW* (Shembade *et al.*, 2008). La proteína A20 regula la señalización mediada por RIP1 a través de inactivación y corte de cadenas de poli-ubiquitina unidas por sus lisinas 48, que conduce a la degradación mediada por proteosoma. Así, Tax1BP1 promueve la reducción de ubiquitinación de RIP1 dependiente de A20 (Iha *et al.*, 2008).

Una función similar que reduce la ubiquitinación de TRAF6 sería llevada a cabo también por la proteína A20, que gracias a su actividad desubiquitinasa, elimina las ubiquitinas con ayuda de la proteína adaptadora Tax1BP1 (Iha *et al.*, 2008).

Por lo tanto, Tax1BP1 inhibe la poliubiquitinación de RIP1 y de TRAF6 al reclutar la proteína A20, que tiene capacidad de añadir o eliminar ubiquitinas y, a través de su unión con la ligasa de ubiquitina Itch, actuar sobre estas moléculas y otros factores para regular la activación de NFκB (Iha *et al.*, 2008).

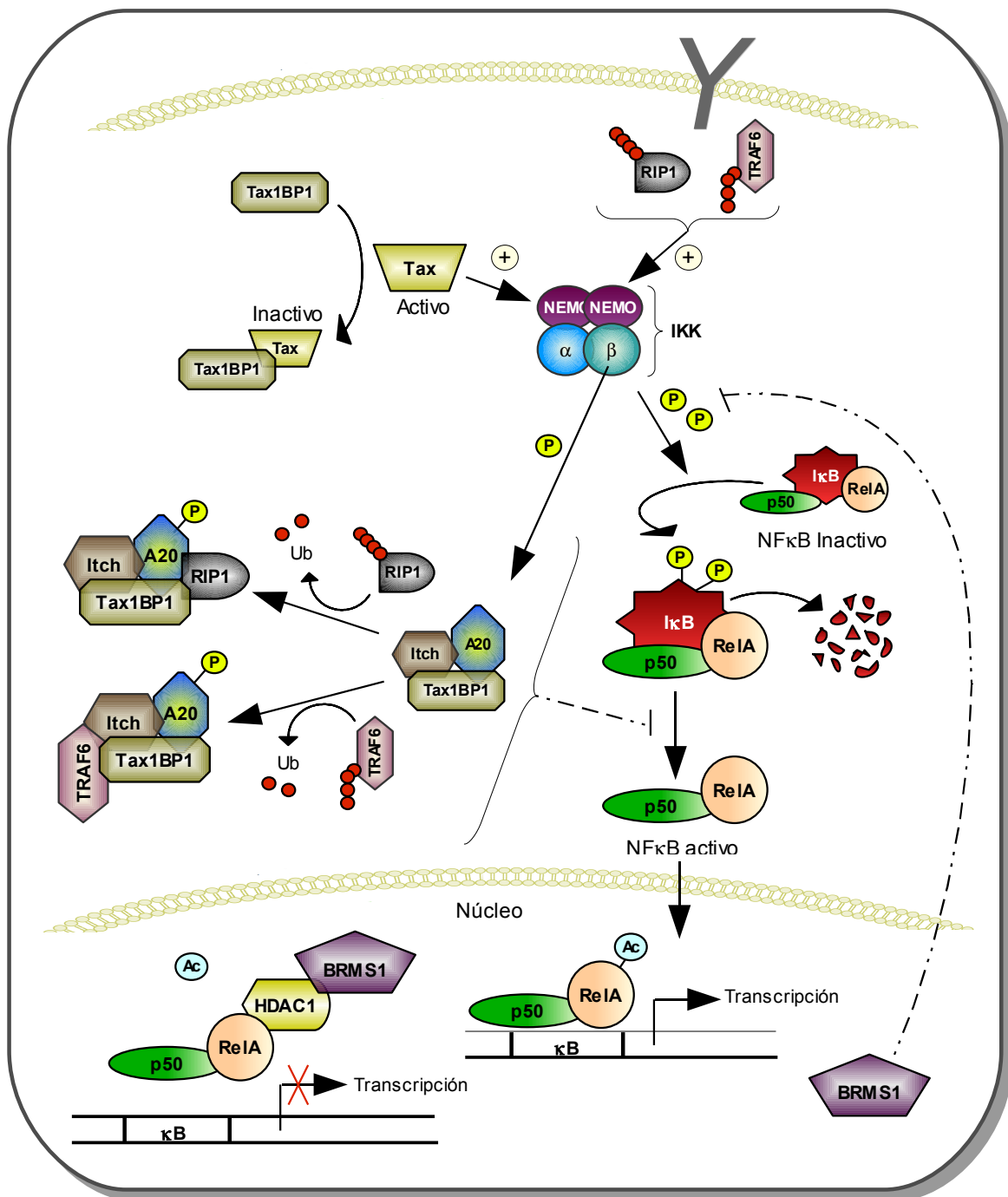


Figura 7. Representación de la regulación inhibitoria de BRMS1 y Tax1BP1 sobre diferentes etapas de la ruta de activación de NFκB.

Por un lado, se muestra cómo Tax1BP1 actúa como un adaptador entre la ligasa de ubiquitina *Itch* y la proteína *A20*, la cual tiene actividad desubiquitinasa dirigida hacia las lisinas 48 y 63 unidas en las cadenas de poliubiquitina. Todo ello forma un complejo multiproteico *A20-Itch-Tax1BP1*, que inhibe la poliubiquitinación de *RIP1* y de *TRAF6*, reprimiendo la ruta de NFκB. También se muestra como Tax1BP1 recluta a *Tax*, inhibiendo su capacidad de activación sobre la ruta (Iha *et al.*, 2008; Shembade *et al.*, 2008).

Por otro lado, el supresor de metástasis BRMS1, ejerce una acción inhibitoria sobre la ruta de dos maneras: inhibe la fosforilación de la subunidad IκBα, impidiendo por lo tanto la activación de NFκB (Cicek *et al.*, 2005; Samant *et al.*, 2007) y, junto con HDAC1, promueve la desacetilación de la subunidad RelA ó p65 (Liu *et al.*, 2006), impidiendo su unión al promotor.

Los grupos fosforilados están representados como "P" en amarillo, y "Ub" se refiere a moléculas de ubiquitina. Cuando el efecto sobre la ruta es de activación, se representa como "+", y cuando es inhibitorio, como "-".

Objetivos

El proceso de metástasis, que representa una de las mayores causas de fallecimiento a nivel mundial, es un complicado mecanismo desencadenado por la desregulación de importantes rutas de señalización celular. Una de ellas es la vía de activación del factor de transcripción NFκB, que controla la regulación de la transcripción génica de proteínas esenciales para las funciones celulares, regulando muchos procesos, como inmunidad, inflamación y apoptosis (Barkett y Gilmore, 1999). Su sobre-activación se correlaciona con cáncer (Karin, 2006; Lin y Karin, 2003). Por ello, resulta indispensable profundizar en el conocimiento de los mecanismos involucrados en esta ruta, así como de las moléculas que ejercen control tanto positivo como negativo sobre la misma.

Las proteínas BRMS1 y Tax1BP1 desempeñan un papel crucial en el contexto de la regulación negativa de la ruta de activación de NFκB (Cicek *et al.*, 2005; Iha *et al.*, 2008). Aunque existe un gran desconocimiento, tanto a nivel biológico como estructural, sobre el supresor de metástasis BRMS1, las evidencias obtenidas previamente en el laboratorio, nos han orientado hacia la selección de relevantes moléculas interactuantes, como es el co-factor de la transcripción NMI. Estos estudios, además de otros descritos en la literatura, que confirman la importancia de los motivos “coiled-coil” en interacción proteína-proteína (Lupas, 1996; Mason y Arndt, 2004), sugieren que la unión entre ambas proteínas está mediada por los mencionados motivos predichos. El conocimiento a nivel estructural, tanto de la proteína BRMS1 como del complejo con NMI, podría facilitar la comprensión de su papel en la vía de NFκB y de manera mas general, en la inhibición de la metástasis.

Por otro lado, el estudio de los mecanismos de señalización sobre diferentes proteínas, bien agonistas o antagonistas de dicha ruta, podría ofrecer una información privilegiada sobre el modo de controlarla. Uno de estos procesos es la ubiquitinación (Chen, 2005), con la que la proteína Tax1BP1 está estrechamente relacionada, y en concreto, a través de sus dos dominios “dedos de zinc” predichos de unión a ubiquitina (UBZs) (Iha *et al.*, 2008). Las interacciones, entre otras, con moléculas de ubiquitina, o con la ligasa de ubiquitina Itch, desempeñan un papel crucial en la regulación directa de la transcripción de genes por parte de NFκB. La caracterización de estos complejos nos ayudaría a profundizar en el conocimiento de sus mecanismos.

Los objetivos a largo plazo del proyecto en el cual se enmarca este trabajo incluyen el establecimiento de nuevas terapias contra la metástasis. Dentro de esta estrategia se precisa tener un conocimiento lo mas ajustado posible de los aspectos moleculares que gobiernan la función de las proteínas BRMS1 y Tax1BP1, para poder así establecer en un futuro las bases estructurales en el desarrollo de fármacos que sean capaces de modular la acción de dichos genes.

Con todos estos indicios, se establecieron los siguiente objetivos en concreto:

1. Determinar la estructura tridimensional de las proteínas BRMS1 y Tax1BP1 y/o de los dominios relevantes para ejercer sus funciones de inhibición sobre la vía de activación de NFκB. Estos dominios son respectivamente: el motivo “coiled-coil” y los dominios “dedos de zinc” predichos de unión a ubiquitina. Se ha elegido la cristalografía de difracción de rayos X como método para alcanzar tal objetivo.

2. Analizar las diferencias estructurales entre los dos dominios *UBZ* de la proteína Tax1BP1 y su implicación en la unión a ubiquitina, realizando una comparación secuencial y estructural de otros dominios similares que interaccionan con ubiquitina ya descritos.
3. Caracterizar biofísicamente las interacciones mediadas entre: BRMS1 y NMI por un lado, y Tax1BP1 con Ubiquitina y con Itch por otro.
4. Determinar la conformación tridimensional de las proteínas BRMS1 y Tax1BP1 en complejo con NMI y con Ubiquitina e Itch respectivamente.

Materiales y Métodos

Para la selección y diseño de las proteínas objeto de interés, se realizó una búsqueda de la información existente de cada una de ellas en diversas bases de datos: *ExPASy*³, *UniProtKB/Swiss-Prot*⁴, *PROSITE*², *BMERC*⁵, *Bioinformatic Harvester (EMBL)*⁶, *NCBI*⁷, *GeneCards*⁸, *Human Protein Reference Database (hprd)*⁹, *PDB*¹⁰ o *PSIPRED*¹¹. Con estas herramientas informáticas se estudiaron las características biológicas, funcionales, físico-químicas y estructurales conocidas de dichas proteínas humanas. En el estudio de la conservación génica entre las diferentes especies se usaron los programas *HomoloGene (NCBI)*¹², *GeneCard*⁸ y *MultAlin*¹³ para realizar alineamientos de secuencias, tanto nucleotídica como aminoacídica.

1. Clonaje de los diferentes amplicones en vectores de expresión procariota

Para el diseño de las diferentes construcciones de BRMS1, NMI y Tax1BP1 humanas con las que se trabajó durante este estudio, se utilizó la información existente en diferentes bases de datos (referencias 1-11), además de otras herramientas informáticas de predicciones de estructura secundaria (*PSIPRED*¹¹) y de motivos “coiled-coil” (como *Coiled-coil prediction*¹⁴, *Multicoil*¹⁵ o *pbil*¹⁶).

El diseño de construcciones para la proteína Tax1BP1 humana, se realizó centrándonos en la predicción de los dos denominados dominios de unión a ubiquitina, mostrados en la Figura 4.

Se diseñaron y clonaron un elevado número de construcciones para las proteínas BRMS1 y NMI humanas (simplificadas en la Tabla VII) y para Tax1BP1.

1.1 Generación de amplicones y construcción de plásmidos recombinantes

La amplificación de secuencias de ADN a clonar se realizó mediante reacción en cadena de la polimerasa (PCR) usando el termociclador “MiniCycler™” (*MJ Research*⁴).

En todos los casos se utilizaron 2,5 unidades de enzima ADN-polimerasa Pfu (*Promega*^B), en el tampón suministrado por la casa comercial (20 mM Tris-HCl pH 8,8; 2 mM MgSO₄; 10 mM KCl; 10 mM (NH₄)₂SO₄; 0,1 % Tritón X-100; 0,1 mg/mL SAB) y complementado con 0,25 mM de una mezcla de desoxinucleótidos trifosfatos (*Fermentas*^C); 0,5 µM de cada uno de los cebadores o iniciadores, que contenían dianas de restricción para facilitar el clonaje, suministrados por *Proligo*^D y 50 ng del ADN molde, en todos los casos de origen humano (Tabla II).

Se realizaron 30 ciclos de amplificación en la que se varió la temperatura de hibridación o anillamiento según las indicadas en la Tabla II, fijándose en 5-8 °C más baja que la temperatura de fusión o disociación de las hebras (Innis and Gelfand, 1990; referencia 1*). La temperatura de fusión o disociación (Td), se calculó siguiendo la fórmula para oligonucleótidos de mas de 20 pb (Breslauer *et al.*, 1986), (herramienta facilitada en la dirección informática de *Proligo*¹⁷).

En el programa de PCR se fijaron las siguiente etapas:

- 5 minutos a 95°C.
- 30 ciclos, cada uno de los cuales con tres etapas. La primera de desnaturalización (30 segundos a 95°C), seguido de una fase de hibridación (1 minuto a la temperatura de anillamiento de los cebadores) y finalmente una etapa de elongación (durante un tiempo determinado según la longitud del fragmento a amplificar (aproximadamente 2 minutos/1 Kb) a 72°C.
- 10 minutos a 72°C.

Todos los fragmentos de ADN procedentes de PCR se purificaron mediante extracción del gel de agarosa. Para tal efecto, se utilizó el “kit” de extracción de gel “QIAquick®” de la casa comercial *Qiagen*^G. Una vez limpios, los productos de PCR, así como los vectores correspondientes para sistema procariota, se digirieron a 37 °C con las indicadas enzimas de restricción (Tabla II) siguiendo las instrucciones del proveedor (*Fermentas*^C).

Los vectores usados procedieron de *Novagen*^H (pET21a) y de *GE Healthcare-Amersham Biosciences*^I (pGEX6P2). El plásmido “pET28-SMT3” se trata de un vector derivado de pET28, con una inserción tras las 6 histidinas en la región amino-terminal del sitio de clonaje múltiple, de una secuencia que codifica para la proteína SMT3 (de unos 100 residuos de longitud) con un sitio de reconocimiento de la proteasa ULP que elimina His-SMT3. Este vector fue cedido por el Dr. Lima (Mossessova y Lima, 2000). El vector abreviado como “pET28-PP”, deriva de una modificación realizada sobre el vector pET28, con adición de un sitio de corte para la proteasa “PreScission” entre las histidinas y la secuencia clonada. pET28-PP fue cedido por el Dr. Ramón Campos (Campos-Olivas *et al.*, 2007).

Para evitar su re-ligación, los vectores se desfosforilaron usando la fosfatasa alcalina con su correspondiente tampón y siguiendo el protocolo establecido por la casa comercial (*Roche*^J).

Tras la limpieza, tanto de los insertos digeridos como de los vectores digeridos y desfosforilados, usando un “kit” para purificación de productos de PCR “JETquick” (*Genomed*^K), se realizaron las reacciones de ligación. Estas reacciones, a 22°C de temperatura, se prolongaron durante 14 horas, usando la enzima ADN ligasa T4 y siguiendo el protocolo de la casa comercial (*Fermentas*^C). Se probaron diferentes reacciones conteniendo tres proporciones molares distintas entre vector:inserto (1:1, 1:3 y 1:5) en un volumen final de 15 µL.

La mezcla de ligación se transformó en células de la cepa DH5α de *E.coli* (*Invitrogen*^L), para lo que se mezcló todo el volumen de ligación (15 µL) con 60 µL de células química-competentes preparadas previamente en el laboratorio siguiendo el protocolo establecido por el autor *Ausubel* (ver referencia 6*). Los viales se mantuvieron a 4 °C durante 30 minutos. La transformación se realizó bajo un choque térmico de la mezcla a 42 °C, durante 90 segundos, y posterior enfriamiento a 4 °C durante 5 minutos. A continuación se añadió 1 mL de medio SOC (8,55 mM NaCl; 20 g/L Triptona; 5 g/L Extracto de levadura; 0,25 mM KCl; 2 M MgCl₂; 1 M Glucosa; pH 7), y se mantuvo una agitación continuada a 500 r.p.m., a 37°C durante 2 horas. Transcurrido este tiempo, se plaqueó todo el volumen de ligación en placas de LB-agar con el antibiótico específico

para cada vector (35 µg/mL kanamicina en el caso de pET28 y 50 µg/ml ampicilina para los vectores pET21a y pGEX) y se incubó a 37 °C toda la noche.

Algunas de las colonias fueron seleccionadas al azar y se cultivaron en medio líquido (LB complementado con el antibiótico) a 37 °C. A continuación se extrajo el ADN plasmídico, para lo que se utilizó el “kit WIZARD® Plus - SV Minipreps” de *Promega*^N, siguiendo las instrucciones dictadas por el proveedor. Los plásmidos se eluyeron y mantuvieron en agua filtrada y estéril.

Para comprobar la correcta inserción del fragmento de ADN de interés, se realizaron sobre el ADN aislado, tanto PCR (de igual manera que la explicada en la sección de “*Generación de amplicones*”) como digestión con enzimas de restricción. Los productos obtenidos de estas dos técnicas se visualizaron por electroforesis de ADN.

Las secuencias de los diferentes insertos de estos plásmidos seleccionados, se validaron mediante secuenciación del ADN.

1.2 Electroforesis de ADN

Las muestras de ADN se analizaron mediante electroforesis en gel de agarosa, a una concentración del 1 % (p/v) en tampón 1xTBE (90 mM Tris-Borato; 2 mM EDTA pH 8), conteniendo colorante fluorescente bromuro de etidio, a una concentración de 0,5 µg/mL. Las electroforesis se realizaron usando una unidad horizontal “PowerPac BasicTM” de *BioRad*^E en tampón 1xTBE.

Las muestras a analizar se prepararon en un tampón de carga conteniendo 0,1 % SDS, 5 % glicerol y 0,05 g/L de azul de bromofenol (*Takara*^F), para facilitar la entrada de las muestras en el pocillo y la visualización en su migración hacia el ánodo.

Tabla II. Materiales usados en los clonajes

Gen	ADN molde	Fragmento aminoácido clonado (aa)	Cebadores	*Ta	Vector de expresión (Proteasa)	Enzimas de restricción	Etiqueta	Construcción aminoácida
BRMS1 (Breast cancer metastasis suppressor 1)	GenBank: BC009834 Prot. ID: AAH09834 Swiss-Prot: Q9HCU9 Clon - ID (rzpd): IRALp962L0425Q2	FL (1-246)	1: 5' GC- <u>agatct</u> -ATGCCTGTCCAGCCTCCAAG3' 2: 5' CG- <u>gtcgac</u> -aagctt-TCAAGGTCCATCCGAT3'	55 °C	pET28-SMT3 (proteasa "ULP")	[BamHI/BglI] / HindIII	6xHis-SMT3 en N-t	a: MISS-HHHHHH-SMT3-Sitio Proteasa* S-BRMS1_{FL} b: S-BRMS1 _{FL}
		51-98	1: 5' GCG-ggatcc-GAGGACTATGAGCGACGCC3' 2: 5' CGC-ctcgag-TTAGGCTCTCTCAGCCCC3'	55 °C	pET28-SMT3 (proteasa "ULP")	BamHI / XhoI	6xHis-SMT3 en N-t	a: MISS-HHHHHH-SMT3-Sitio Proteasa* S-BRMS1₅₁₋₉₈ b: S-BRMS1 ₅₁₋₉₈
		51-84	1: 5' GCG-ggatcc-GAGGACTATGAGCGACGCC3' 2: 5' GC-ctcgag-TTAACTCAGTCGTTCCCTGAAC3'	50 °C	pET28-SMT3 (proteasa "ULP")	BamHI / XhoI	6xHis-SMT3 en N-t	a: MISS-HHHHHH-SMT3-Sitio Proteasa* S-BRMS1₅₁₋₈₄ b: S-BRMS1 ₅₁₋₈₄
NMI (N-Myc interactor)	GenBank: NM_004688 Prot. ID: NP_004679 Swiss-Prot: Q13287 Clon: Librería cDNA glándula mamaria	FL (1-307)	1: 5' GC-catatg-GAAGCTGATAAAGATGACAC3' 2: 5' CG-ctcgag-CTATTCTCAAAGTATGC3'	50 °C	pET21a	NdeI / XhoI	-	NMI_{FL} -stop
		16-64	1: 5' GCG-ggatcc-CCAGATGAATTTATAAAAGAT3' 2: 5' GC-ctcgag-TTAAATATCCTCTTTAATCTGG3'	50 °C	pET28-SMT3 (proteasa "ULP")	BamHI / XhoI	6xHis-SMT3 en N-t	a: MISS-HHHHHH-SMT3-Sitio Proteasa* S-NMI₁₆₋₆₄ b: S-NMI ₁₆₋₆₄
Tax1BP1 (Tax1 (human T-cell leukemia virus type 1) binding protein 1)	GenBank: NM_006024 Prot. ID: NP_006015 Swiss-Prot: Q86VP1 Clon proporcionado por el Dr. Ivan Dikic (Iha et al., 2008)	UBZ1 + UBZ2 (725-789)aa	1: 5' GCG-catatg-GATGTTCAACAAGAGTGTC33' 2: 5' GC-ctcgag-CTAGTCAAAATTTAGAACATTC3'	55 °C	pET28-PP (Proteasa "PreScission")	NdeI / XhoI	6xHis en N-t	a: MGSS-HHHHHH-SSG- <u>LEVLFQ</u> *GP-HM- UBZ1+2 b: GPHM- UBZ1+UBZ2
Itch (Itchy homolog E3 ubiquitin protein ligase)	GenBank: NM_031483 Prot. ID: NP_113671 Swiss-Prot: Q96J02 Clon - ID (CNIO): hAF3512	WW ₃ +WW ₄ (435-511)aa	1: 5' GCG-ggatcc-GATCCTCTTGGTCCATTGCC3' 2: 5' GC-ctcgag-TTATGTGCGGGGATCTATATA3'	55 °C	pGEX6P2 (Proteasa "PreScission")	BamHI / XhoI	GST en N-t	a: GST- <u>LEVLFQ</u> *GP-LGS- WW₃+WW₄ b: GPLGS- WW₃+WW₄

Se muestran los materiales usados para los clonajes de las construcciones representadas según el fragmento aminoácido para el que codifican. En la segunda columna se exponen los códigos identificativos de las bases de datos de los genes y su correspondiente proteína, así como la fuente del clon usado.

1: Cebador en el extremo 5' (iniciador hacia adelante) - **2:** Cebador en el extremo 3' (iniciador reverso) / ***Ta** = Temperatura de anillamiento usada en la reacción de amplificación / **a:** Secuencia aminoácida antes de proteólisis con la proteasa específica - **b:** Secuencia aminoácida final tras la proteólisis con la proteasa específica / **Subrayado:** Secuencia de reconocimiento de la proteasa específica para cada vector / *: Sitio de corte de la proteasa. La procedencia de los vectores usados se explica en la sección "1.2.3-Digestión de los vectores y amplicones con enzimas de restricción" / **aa:** aminoácidos. Las enzimas de restricción introducidas en corchetes ([]) son compatibles, y generaron una nueva diana.

UBZ: dominio estructurado en "dedo de zinc" y predicho de unión a ubiquitina (Iha et al., 2008). / **WW:** motivos conteniendo dos triptófanos conservados y de unión a dominios PPxY (Shembade et al., 2008; Sudol et al., 1995).

2. Expresión de proteínas recombinantes

Los plásmidos recombinantes generados fueron transformados mediante electroporación (usando un voltaje de 1,8 kV). Para ello se mezclaron 50 ng de ADN con 60 µl de diferentes cepas de expresión de *E.coli* electrocompetentes, en función de las características de la proteína a expresar, como son BL21(DE3) y Rosetta(DE3)pLysS, preparadas previamente en el laboratorio según el protocolo descrito por el autor Ausubel (ver referencia 6*), y C41(DE3) y TG1tr (*Invitrogen*^L).

Tras el choque eléctrico, las células se recuperaron con 1 mL de medio SOC, a 37°C durante 1 hora. Las células transformadas se seleccionaron en placas de LB-agar complementado con el antibiótico específico para cada vector. Se eligieron varias colonias al azar para realizar los ensayos de expresión.

Para la co-expresión de las secuencias codificantes completas de BRMS1 y NMI (ver Tabla II), se realizó la co-transformación de ambos vectores, siguiendo el protocolo descrito anteriormente. La selección de colonias co-transformadas se realizó en placas de LB-agar conteniendo 35 µg/mL kanamicina y 50 µg/mL ampicilina, para pET28-SMT3-BRMS1 y pET21-NMI respectivamente.

Los ensayos de sobre-expresión de la proteína se realizaron en diferentes medios de crecimiento: LB o “*Luria-Bertoni*” (0,17 M NaCl; 1 % Triptona; 0,5 % Extracto de levadura; pH 7,2) y 2xTY o “*2 times Tryptone-Yeast*” (85 mM NaCl; 1,6 % Triptona; 1 % Extracto de levadura; pH 7,2), complementados con antibiótico (35 µg/mL kanamicina en el caso de células conteniendo vectores de expresión derivados de pET28 y 50 µg/ml ampicilina para los vectores pET21a y pGEX).

Se crecieron cultivos iniciadores con 5-10 mL de volumen a 37 °C durante 14-16 horas, y posteriormente se diluyeron 1:20 con medio fresco y se crecieron hasta alcanzar una densidad óptica de 0,6; tras lo que se probaron diferentes concentraciones (0,1 mM, 0,5 mM y 1 mM) del inductor IPTG (*Isopropil-β-D-tiogalactopiranosido*), así como diferentes temperaturas (15 °C, 20 °C, 25 °C, 30 °C y 37 °C). El tiempo de expresión fue distinto según la temperatura, de forma que para ensayos a 15 °C, 20 °C y 25 °C la inducción de la expresión se mantuvo durante toda la noche y para 30 °C y 37 °C se probaron diferentes intervalos de tiempo (transcurridos 2, 4 y 6 horas).

Una vez optimizadas las condiciones de expresión para cada proteína, se crecieron cultivos celulares a gran escala, usando volúmenes de 1 L de medio de crecimiento en matraces truncados de 2,8 L de capacidad, manteniendo una agitación de 250 r.p.m.

Transcurrido el tiempo óptimo de expresión de añadir el IPTG, la masa celular resultante de estos ensayos de expresión fue usada para la extracción de las proteínas sobre-expresadas (purificación).

2.1 Expresión de proteínas derivadas con selenio-metionina (SeMet - BRMS1₅₁₋₉₈)

El plásmido pET28-SMT3-BRMS1₅₁₋₉₈ se transformó en células de expresión no auxotróficas *Rosetta(DE3)*, mediante electroporación, y posterior selección en placas de LB-agar conteniendo 35 µg/mL de kanamicina.

A continuación se inoculó la colonia seleccionada en 50 mL de LB complementado con el antibiótico y se creció durante toda la noche a 37 °C. Este cultivo iniciador fue diluido en 1 L de LB y crecido hasta alcanzar una densidad óptica a 600 nm de 0,6. A continuación el cultivo se centrifugó durante 10 minutos, a 6.000 x g y a 4 °C. La masa celular resultante se resuspendió en 30 mL de medio M9 (componentes expuestos al final del apartado) para eliminar el LB residual y se centrifugó a 1.200 x g. El sedimento se resuspendió en 1 L de medio M9 enriquecido y se dejaron recuperar durante 30 minutos a 37 °C en agitación. Tras este tiempo, se añadió el cóctel de aminoácidos directamente en polvo sobre el medio de cultivo, dejando crecer 15 minutos y añadiendo 1 mM de IPTG. La inducción se mantuvo durante 4 horas a 37 °C en agitación.

Composición de medios (autoclavar las soluciones y filtrar la vitamina y glucosa, con filtros de 0,2 µm):

- *Medio M9 enriquecido (1X)*: Medio M9_1X; Solución Ca/Mg; 10 mg/L Vitamina B1 y 0,4 % Glucosa.
- *Medio M9 (1X)*: 8,8 g/L Na₂HPO₄-2H₂O; 3 g/L KH₂PO₄; 0,5 g/L NaCl y 1 g/L NH₄Cl.
- *Solución Ca/Mg (para 1 L de cultivo)*: 14,7 mg/L CaCl₂-2H₂O; 246 mg/L MgSO₄-7H₂O.
- *Cóctel de aminoácidos (para 1 L de cultivo)*: 50 mg SeMet; 100 mg Lisina; 100 mg Treonina; 100 mg Fenilalanina; 50 mg Leucina; 50 mg Isoleucina y 50 mg Valina.

2.2 Expresión de proteínas marcadas con Nitrógeno 15 (¹⁵N-BRMS1₅₁₋₈₄)

Se inoculó la colonia seleccionada portadora del plásmido recombinante *pET28-SMT3-BRMS1₅₁₋₈₄* en medio LB complementado con 35 µg/mL de kanamicina y se procedió de la misma manera que para proteínas no derivadas, hasta obtener un cultivo celular, que se centrifugó a 6.000 x g, durante 10 minutos y a 4 °C de temperatura, seguidos de dos lavados consecutivos del sedimento con el medio M9 (ver abajo composición de medios), el primero con 800 mL y el segundo con 500 mL. Finalmente, la masa resultante se resuspendió en 1 L de medio mínimo completo complementado con ¹⁵N y antibiótico, y se creció durante 1 hora a 37 °C. Por último se añadió 0,5 mM IPTG, manteniéndose durante 4 horas a 37 °C, en agitación.

Composición de medios (autoclavar el medio de sales M9 y filtrar el resto de las soluciones con filtro de 0,2 μm):

- *Medio mínimo completo marcado (1X)*: Medio de sales M9_1X; Elementos Trace; 0,3 mM CaCl_2 ; 1 mM MgSO_4 ; 1 mg/L D-Biotina; 30 mg/L Tiamina y 1 mg/L extracto de levadura.
Para ser añadido en el último momento: 1 gr/L NH_4Cl (^{15}N); 2 g/L Glucosa.
- *Medio de sales M9 (1X)*: 13 g/L KH_2PO_4 ; 10 g/L K_2HPO_4 ; 9 g/L Na_2HPO_4 ; 2,4 g/L K_2SO_4 ; ajustar a pH 7,2 con NaOH sólido.
- *Elementos Trace (para 1 L de cultivo)*: 6 mg/L $\text{FeSO}_4\cdot 7\text{H}_2\text{O}$; 6 mg/L $\text{CaCl}_2\cdot 2\text{H}_2\text{O}$; 1,2 mg/L $\text{MnCl}_2\cdot 4\text{H}_2\text{O}$; 0,8 mg/L $\text{CoCl}_2\cdot 6\text{H}_2\text{O}$; 0,7 mg/L $\text{ZnSO}_4\cdot 7\text{H}_2\text{O}$; 0,3 mg/L $\text{CuCl}_2\cdot 2\text{H}_2\text{O}$; 0,02 mg/L H_3BO_3 ; 0,25 mg/L $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24}\cdot 4\text{H}_2\text{O}$ y 5 mg/L $\text{EDTA}\cdot \text{Na}_2(\text{H}_2\text{O})_2$.

3. Purificación de proteínas

Tras la sobre-expresión, los cultivos celulares (2-3 litros de volumen usados en cada purificación) se centrifugaron a 6.000 x g durante 10 minutos, a 4 °C de temperatura. La masa celular se resuspendió en 35 mL de tampón de lisis por litro de cultivo (ver tampón específico para cada construcción en las Tablas III-A, para las construcciones unidas a histidinas y IV-A para aquellas fusionadas a GST), complementado con inhibidores de serín- y cisteín-proteasas “Complete, EDTA-free” de la casa comercial *Roche*¹, siguiendo para su uso, el protocolo de la casa comercial.

Una vez fueron lisadas las células mediante sonicación (sonicador “Vibracell 75042” de *Fisher Scientific Bioblock*^R), se centrifugó durante 50 minutos a alta velocidad (125.000 x g) para separar las fracciones celulares solubles de las insolubles. Tanto la sonicación como la centrifugación se realizaron a una temperatura de 4 °C.

En los casos donde se requirió mejorar la solubilidad de la proteína tras la lisis bacteriana, se probaron diferentes agentes químicos en el tampón de lisis, como son: 0,5 % Tween-20; 0,5 % Tritón X-100 o 10 % glicerol. Los tampones finales usados para la lisis celular se muestran en las Tablas III-A y IV-A.

La muestra filtrada (0,45 μm de diámetro de poro) se cargó en las columnas indicadas en las siguientes secciones, con ayuda de una bomba peristáltica a una velocidad de entre 2 y 5 mL/min, según la proteína (2 mL/min para proteínas de fusión a GST y de 3 a 5 mL/min para proteínas unidas a histidinas).

Las purificaciones se realizaron usando un *ÄKTApurifier*TM (para las columnas de afinidad) y un *ÄKTAprime*TM (para las columnas de filtración en gel), ambos de *GE Healthcare-Amersham Biosciences*¹. La monitorización y seguimiento de las proteínas en cada etapa de la purificación se realizó visualizando un cromatograma de absorción de luz ultravioleta a 280 nm de longitud de onda facilitado por el programa “Unicorn 4.00” para las columnas de afinidad y “PrimeView / PrimeView Evaluation” para las de filtración en gel.

Para llevar a cabo los diferentes pasos de las purificaciones, se siguieron los protocolos previamente establecidos en el laboratorio, tratándose de igual manera tanto las proteínas no derivadas como las derivadas de SeMet o marcadas con N¹⁵. Estos protocolos se detallan a continuación:

3.1 Purificación de las construcciones con cola de histidinas

- His-BRMS1₅₁₋₈₄, His-BRMS1₅₁₋₉₈, *SeMet*-His-BRMS1₅₁₋₉₈
- His-NMI₁₆₋₆₄
- His-BRMS1_{FL} + NMI_{FL} (*Co-purificación*)
- His-Tax1BP1₇₂₅₋₇₈₉ (*UBZ1+UBZ2*), His-Tax1BP1₇₂₅₋₇₅₆ (*UBZ1*), His-Tax1BP1₇₅₅₋₇₈₉ (*UBZ2*)

El sobrenadante filtrado fue cargado en una columna de afinidad (“HiTrap Chelating HP” de 5 mL, de *GE Healthcare-Amersham Biosciences*[†]) a la velocidad descrita anteriormente. Dicha columna fue cargada previamente con 0,1 M de sulfato de níquel, siguiendo las instrucciones del fabricante.

Las proteínas de fusión a histidinas retenidas en la columna, fueron posteriormente lavadas con 10 volúmenes de columna de la disolución de lavado. A continuación, la proteína fue eluída de la columna usando un gradiente continuo de imidazol (hasta 0,5 M) en el tampón de lavado durante 13 volúmenes de columna, y a una velocidad de 3 mL/min. La elución se recogió en fracciones de 4 mL cada una. Los tampones de lavado y elución para cada construcción están especificados en la Tabla III-B y C.

Las fracciones de interés fueron tratadas con la proteasa específica indicada para cada construcción en la Tabla II, a una concentración 1:1000 (peso proteasa:peso proteína de fusión), dializando frente a un volumen de tampón de diálisis al menos 500 veces mayor que el de la muestra (Tabla III-D). Para la diálisis se usaron “Slide-A-Lyzer Dialysis Cassettes” de 3,5 kDa de tamaño molecular de diámetro de poro (*Pierce*[§]). La proteólisis y diálisis se realizaron de forma diferente para cada enzima: en dos pasos para la proteasa “ULP”; primero en una solución inicial durante 2 h, y posteriormente cambiando a un nuevo tampón durante toda la noche a 4 °C y para la proteasa “PreScission” durante 90 minutos a temperatura ambiente.

Las proteasas usadas para la eliminación de la cola de histidinas fueron obtenidas en el laboratorio mediante expresión en células de *E.coli* de un vector conteniendo el ADN codificante para la enzima. Para el caso de la proteasa “ULP”, el vector de expresión es pET28a y para la proteasa “PreScission”, pGEX4T1. Posteriormente, estas dos proteasas fueron purificadas de acuerdo a protocolos establecidos en el laboratorio.

Tras la proteólisis, la eliminación de las proteínas de fusión no digeridas y los fragmentos fusionados a histidinas, en lo que respecta a las purificaciones de las construcciones de BRMS1 y NMI clonadas en el vector pET28-SMT3, se realizó mediante una segunda columna de afinidad (“HiTrap Chelating HP” de 5 mL, de *GE Healthcare-Amersham Biosciences*[†]), cargada con 0,1 M de sulfato de níquel y equilibrada con el tampón de diálisis. Los tampones de lavado y elución para esta columna se correspondían

con el de diálisis y el de elución de la columna inicial, respectivamente. Para la inyección, lavado y elución de esta columna se usó una bomba peristáltica, a 3 mL/min de velocidad.

La muestra que no se retuvo en la columna (proteína de interés sin histidinas) se concentró en un concentrador “Amicon Ultra” de 5 kDa de diámetro de poro (*Millipore*^T), y fueron inyectadas en una columna de filtración en gel o exclusión por tamaño molecular (“Superdex 75 16/60 ó 26/60”, de *GE Healthcare-Amersham Biosciences*^I, usando un bucle de inyección de 2 mL y 5 mL de capacidad respectivamente), a una temperatura de 4 °C. Los tampones usados en el equilibrado y la elución, están reflejados en la Tabla III-E. Se colectaron fracciones de 4 mL de volumen y a 0,5 mL/min de velocidad y de 5 mL y a 1 mL/min de velocidad para las columna 16/60 y 26/60 respectivamente. La estimación de los tamaños moleculares obtenidos se realizó interpolando los volúmenes de elución en una curva de calibración obtenida previamente en el laboratorio con estándares de tamaño molecular conocidos (*Sigma*^O). Las fracciones de interés se concentraron mediante unidades de ultrafiltración “Amicon Ultra” de 5 kDa de tamaño de poro (*Millipore*^T), a 4 °C para su posterior utilización en diferentes experimentos de caracterización biofísica y cristalográficos.

La co-purificación de las dos secuencias codificantes completas para BRMS1, fusionado a histidinas, y para NMI, se realizó tal y como se expone en este apartado, tratando, a fines prácticos, la mezcla proteica como una sola proteína.

	Columna de afinidad				
	A	B	C	D	E
	Tampón de lisis	Tampón de lavado	Tampón de elución	Tampón de diálisis	Tampón de filtración en gel
Tax1BP1 (UBZ1 + UBZ2)	- 50 mM TRIS - 0,5 M NaCl - 0,5 % (v/v) Tween-20	- 50 mM TRIS - 0,2 M NaCl - 10 mM Imidazol	- 50 mM TRIS - 0,2 M NaCl - 0,5 M Imidazol	- 50 mM TRIS - 0,2 M NaCl	- 20 mM TRIS - 0,2 M NaCl
BRMS1 (51-84)	- 50 mM TRIS - 0,5 M NaCl - 0,5 % (v/v) Tween-20	- 50 mM TRIS - 0,15 M NaCl	- 50 mM TRIS - 0,15 M NaCl - 0,5 M Imidazol	- 50 mM TRIS - 0,15 M NaCl	- 20 mM TRIS - 50 mM NaCl - +/- 2 mM DTT*
BRMS1 (51-98)	- 50 mM TRIS - 0,5 M NaCl - 0,5 % (v/v) Tween-20	- 50 mM TRIS - 0,15 M NaCl - 10 mM Imidazol	- 50 mM TRIS - 0,15 M NaCl - 0,5 M Imidazol	- 50 mM TRIS - 0,15 M NaCl	- 20 mM TRIS - 50 mM NaCl
NMI (16-64)	- 50 mM TRIS - 0,5 M NaCl - 0,5 % (v/v) Tween-20	- 50 mM TRIS - 0,15 M NaCl	- 50 mM TRIS - 0,15 M NaCl - 0,5 M Imidazol	- 50 mM TRIS - 0,2 M NaCl	- 20 mM TRIS - 50 mM NaCl
BRMS1_{FL} + NMI_{FL}	- 50 mM TRIS - 0,5 M NaCl - 0,5 % (v/v) Tween-20	- 50 mM TRIS - 0,15 M NaCl - 10 mM Imidazol	- 50 mM TRIS - 0,15 M NaCl - 0,5 M Imidazol	- 50 mM TRIS - 0,15 M NaCl	- 20 mM TRIS - 50 mM NaCl

Tabla III. Tampones usados en las purificaciones de las construcciones con cola de histidinas, diseñadas para las proteínas Tax1BP1, BRMS1 y NMI. Se exponen las disoluciones comunes usadas para todas las construcciones de la misma proteína. Los tampones “A”, “B” y “C” aluden respectivamente a los usados en los procesos de lisis, lavado y elución de la columna de afinidad; el tampón “D” al usado durante la diálisis-

proteolisis y el denominado con la letra “E” se refiere al de la columna de exclusión por tamaño molecular. Cabe destacar que todos los tampones compuestos por TRIS se ajustaron a un pH de 7,5 a 4 °C.

* Para ensayos de cristalización de la construcción BRMS1 aa51-84 sola, o bien ensayos bioquímicos para caracterizar sus interacciones, no se añadió DTT a la proteína final. Sin embargo, para ensayos de cristalización destinados a la obtención de la estructura de complejos, se añadió 2 mM del agente reductor con el fin de evitar la formación de puentes disulfuro.

3.2 Construcciones fusionadas a GST

- Poli4-Ubiquitina
- Itch₃₂₃₋₃₉₁ (WW₁+WW₂), Itch₄₃₅₋₅₁₁ (WW₃+WW₄), Itch₃₂₃₋₅₁₁ (WW₁-WW₄)

El lisado celular filtrado procedente de la sobre-expresión de las construcciones indicadas, fue inyectado en una columna de afinidad por glutatión S-transferasa (“GSTrap HP” de 5 mL, de *GE Healthcare-Amersham Biosciences*[†]) a una velocidad máxima de 2 mL/min, siguiendo el protocolo de la casa comercial (Tabla IV-A).

Las proteínas de fusión a GST retenidas en la columna, fueron posteriormente lavadas con 5 volúmenes de columna de la disolución de lavado a una velocidad de 2 mL/min. A continuación, la proteína fue eluída de la columna usando 10 volúmenes de columna del tampón de elución (conteniendo 10 mM de glutatión reducido ajustado a un pH de 8,0) a una velocidad de 3 mL/min, a través de fracciones de 5 mL. Los tampones de lavado y elución para cada construcción están especificados en la Tabla IV-A y B.

Para proceder con la digestión enzimática, se añadió la proteasa específica (Tabla II) a las fracciones conteniendo la proteína eluída, a una concentración 1:1000 en peso proteasa:peso proteína de fusión para la proteasa “*PreScission*” (obtenida en el laboratorio tal y como se explicó en el apartado anterior). En el caso de proteolisis con *Trombina* (*GE Healthcare-Amersham Biosciences*[†]), se siguieron las instrucciones del proveedor, por lo que se añadió 1 unidad de proteasa frente a 100 µg de proteína de fusión.

Todas las reacciones de proteolisis se realizaron dializando frente a un volumen de tampón de diálisis (Tabla IV-C) de 500 veces superior al de la muestra, a temperatura ambiente, mediante membranas de 3,5 kDa de tamaño molecular de diámetro de poro (“Slide-A-Lyzer Dialysis Cassettes”, de *Pierce*[§]). La proteolisis y diálisis se realizaron a temperatura ambiente, durante 90 minutos para el caso de la proteasa “*PreScission*” y durante toda la noche para la digestión con *Trombina*.

Sólo para el caso de la purificación de “poli4-Ubiquitina” fue necesario intercalar una segunda columna de afinidad (“GSTrap HP” de 5 mL, de *GE Healthcare-Amersham Biosciences*[†]) para eliminar la GST. Esta columna fué equilibrada con unos 5 volúmenes del tampón de diálisis. Como tampón de lavado se usó el de la diálisis y el tampón de elución fue el mismo que para la columna anterior. Para inyectar la muestra se usó una bomba peristáltica a una velocidad de 2 mL/min y a 3 mL/min para la elución.

La mezcla proteica conteniendo los productos de digestión (muestra resultante tras la proteolisis para el caso de las construcciones de Itch, y la fracción no retenida en la segunda columna de afinidad para el caso de poli4-Ubiquitina) se concentró a 4 °C en un

concentrador “Amicon Ultra” de 5 kDa de diámetro de poro para las construcciones de la proteína Itch y de 10 kDa para la poli-4Ub (*Millipore^T*).

Finalmente, las muestras se inyectaron en una columna de exclusión por tamaño molecular (“Superdex 75 16/60 ó 26/60”, de *GE Healthcare-Amersham Biosciences^I*, usando un bucle de inyección de 2 mL y 5 mL de capacidad respectivamente). La cromatografía se realizó a una temperatura de 4 °C, y los volúmenes y velocidades de elución fueron de 4 mL, a 0,5 mL/min para la columna 16/60 y de 5 mL, a 1 mL/min para la 26/60. Los tampones usados en cada caso para equilibrar y eluir las proteínas separadas por su tamaño molecular están reflejados en la Tabla IV-D.

	Columna de afinidad			
	A	B	C	D
	Tampón de lisis = tampón de lavado	Tampón de elución	Tampón de diálisis	Tampón de filtración en gel
Poli-4Ub	- 50 mM TRIS - 0,15 M NaCl - 0,5 % (v/v) Tween-20	- 50 mM TRIS - 0,2 M NaCl - 10 mM Glutación reducido; pH 8,0	- 50 mM TRIS - 0,2 M NaCl	- 20 mM TRIS - 50 mM NaCl
Itch (WW₃ + WW₄)	- 50 mM TRIS - 0,15 M NaCl - 1 % (v/v) Tritón X-100	- 50 mM TRIS - 0,2 M NaCl - 10 mM Glutación reducido; pH 8,0	- 50 mM TRIS - 0,2 M NaCl	- 20 mM TRIS - 0,2 M NaCl

Tabla IV. Tampones usados en las purificaciones de las construcciones fusionadas a GST, diseñadas para las proteínas poli4-Ubiquitina e Itch. Poli-4Ub es la abreviatura utilizada para designar la proteína formada por 4 moléculas de ubiquitina expresadas de forma sucesiva. Se exponen las disoluciones comunes usadas para todas las construcciones de la misma proteína. Los tampones de lisis y lavado (“A”) y de elución (“B”) se usaron en la columna de afinidad a GST, el tampón “C” durante la diálisis-proteólisis y el denominado con la letra “D” se refiere al de la columna de exclusión por tamaño molecular. Cabe mencionar que todos los tampones compuestos por TRIS se ajustaron a un pH de 7,5 a una temperatura de 4 °C, aunque el tampón de elución B se ajustó finalmente a un pH de 8,0.

Las fracciones de interés se concentraron usando concentradores de las mismas características que los usados en la etapa previa a la inyección en la columna de filtración en gel, a 4 °C. Finalmente, para estimar los tamaños moleculares, estos volúmenes de elución se interpolaron en una curva de calibración obtenida con estándares de tamaño molecular conocidos (*Sigma^Q*).

Tras su purificación, las identidades de las proteínas en estado de elevado grado de pureza fueron confirmadas mediante espectroscopía de masas (*MALDI-TOF*) por el servicio de proteómica, en el mismo centro de realización del proyecto.

4. **Obtención de ubiquitina y poli4-Ubiquitina humanas**

La ubiquitina recombinante humana (con 76 aminoácidos y un peso molecular de 8,5 kDa) usada fue adquirida a una elevada pureza y en estado liofilizado, de la casa comercial *BostonBiochem*^N (código U-100H). Para su dilución se utilizó un tampón conteniendo 20 mM TRIS pH 7,5 y 0,2 M NaCl.

El vector de expresión pGEX4T1 portando la secuencia codificante para la Poli-4Ub (cuatro moléculas de ubiquitinas clonadas de forma continuada una tras otra sin dejar residuos entre ellas) fué proporcionado por el Dr. Dikic (Iha *et al.*, 2008). Este vector expresa la proteína unida a GST en su extremo amino terminal seguida de un sitio de reconocimiento de *Trombina*, que tras su proteolisis, genera una proteína sin GST y con 6 aminoácidos extras (GSPGIH-Ub-Ub-Ub-Ub).

5. **Análisis físico-químicos y cuantificación de proteínas**

5.1 **Cuantificación de proteínas**

Los dos principales métodos empleados para la determinación de la concentración de proteínas fueron:

5.1.1 **Espectroscopía de luz ultravioleta a 280 nm.**

En ciertos casos se midió la concentración de proteínas usando la absorbancia a una longitud de onda de 280 nm (para las construcciones conteniendo anillos aromáticos triptófano, tirosina y fenilalanina). La concentración fue calculada mediante la ley de “*Lambert-Beer*” (Ecuación 1):

$$A_{280} = \epsilon d C$$

Ecuación 1. Ley de “*Lambert-Beer*” para la estimación de la concentración de proteína. “ A_{280} ” es la absorbancia medida a partir de la muestra, “ ϵ ” es el coeficiente de extinción molar (propiedad física de cada sustancia, calculada con la herramienta informática *ProtParam*¹⁸), “ d ” la distancia en cm recorrida por la luz dentro de la solución (diámetro de la cubeta, que se estandariza a 1 cm), y “ C ” es la concentración de la molécula en cuestión.

En algunos casos, la concentración de proteínas se estimó utilizando el aparato “nanodrop ND-1000” de la casa comercial *NanoDrop*^{AE}, usando un recorrido de la luz a través de la muestra de 0,2 mm.

5.1.2 Método de colorimetría: Bradford

El colorante utilizado fue “Coomassie Blue G-250” de *BioRad*^E. La medición se realizó mezclando en una cubeta 1 mL del reactivo de “Bradford” (preparado según la casa comercial), con un volumen determinado de la muestra, que se mantuvo a temperatura ambiente durante 10 minutos, para luego detectar la absorbancia a 595 nm frente al blanco. La absorbancia resultante se interpoló en una curva patrón realizada con cantidades conocidas de seroalbúmina bovina, obteniendo así la concentración.

5.2 Análisis de los parámetros físico-químicos de cada proteína

Para determinar los parámetros físico-químicos de todas las construcciones estudiadas en este trabajo, se utilizó la herramienta informática *ProtParam*¹⁸ de “Expasy”. Con ello se calcularon los pesos moleculares, los puntos isoeléctricos teóricos y los coeficientes de extinción molar de las proteínas, así como el contenido en aminoácidos, la vida media y el índice de estabilidad de cada una de ellas.

6. Ensayos biofísicos

6.1 Espectroscopía de Dicroísmo Circular (CD)

Los espectros de dicroísmo se realizaron en un espectropolarímetro JASCO J-810, y usando el programa informático *Spectra Manager*. El sistema fue purgado previamente con gas de nitrógeno a un flujo de 15-20 L/min para eliminar el oxígeno.

Las muestras a estudiar se dializaron frente a un volumen 500 veces superior de un tampón conteniendo 10 mM HEPES pH 7,5 y 50 mM NaCl, a 4 °C durante toda la noche. Posteriormente las muestras fueron diluidas a las concentraciones 0,25 y 0,5 mg/mL para BRMS1₅₁₋₈₄ y NMI₁₆₋₆₄ respectivamente (concentración máxima a la que el voltaje no supera 800 V) y finalmente centrifugadas durante 10 minutos a 16.000 x g. Se utilizó una cubeta de cuarzo de 2 mm de paso de luz (110-QS de *HELLMA*®) con 400 µL de muestra y a 20 °C de temperatura.

Se midió la curva de absorción de las distintas conformaciones, obteniendo la elipticidad en unidades de CD (miligrados), en un rango de longitud de onda entre 190-250 nm, a la que se sustrajo la señal procedente del tampón utilizado.

6.2 Dispersión dinámica de la luz (DLS)

La proteína en solución fue analizada en un tampón 20 mM TRIS pH 7,5 y 50 mM NaCl y a una concentración de 1 mg/mL. Las muestras fueron previamente centrifugadas durante 10 minutos a 16.000 x g.

La dispersión de la luz se midió a 20 °C de temperatura, en un sistema “DynaPro” (*PROTEINSOLUTION*), usando una cubeta de cuarzo de 45 µL y 3 mm de paso de luz. Se realizaron 30 adquisiciones, de 5 segundos de duración cada una, a un potencia del haz de luz de 100 %. Los resultados fueron visualizados y analizados con el programa *Dynamics V6*, con el que se determinó la homogeneidad o porcentaje de polidispersidad, el radio de las partículas y el tamaño molecular.

7. Ensayos de caracterización de interacción proteína-proteína

7.1 Calorimetría de titulación isotérmica (ITC)

Las muestras BRMS1₅₁₋₈₄ cc y NMI₁₆₋₆₄ fueron dializadas a 4 °C durante toda la noche, frente al mismo tampón conteniendo 10 mM HEPES pH 7,5 y 50 mM NaCl. Finalmente se ajustaron las concentraciones de ambas proteínas a 950 µM y 95 µM para BRMS1₅₁₋₈₄ (en la jeringa) y NMI₁₆₋₆₄ (en la celda) respectivamente, fueron desgasificadas y analizadas en un microcalorímetro VP-ITC de MicroCal™.

La titulación se llevó a cabo a 20 °C y una agitación de 310 rpm. Se realizó un total de 25 inyecciones de BRMS1₅₁₋₈₄ sobre NMI₁₆₋₆₄, cada una de ellas de un volumen de 10 µL y una duración de 20 segundos. El tiempo de recuperación entre las inyecciones fue de 120 segundos. La potencia de referencia se fijó en 11,5 µcal/seg y el período de filtro de 2 segundos.

Los resultados fueron visualizados y analizados con *VPViewer2000* y *Microcal, LIC, ITC* respectivamente, del paquete *Origin7*, donde se calcularon los parámetros estequiométricos de la reacción y el tipo de reacción. Para ello se sustrajo la respuesta calorimétrica de dos ensayos: tampón inyectado sobre NMI₁₆₋₆₄ y BRMS1₅₁₋₈₄ inyectado sobre tampón.

7.2 Resonancia Magnética Nuclear (RMN)

Las muestras BRMS1₅₁₋₈₄ cc + NMI₁₆₋₆₄ se analizaron mediante RMN bidimensional. Para ello fueron expresadas, purificadas y concentradas hasta 300 µM. BRMS1₅₁₋₈₄ fue marcada con ¹⁵N según el protocolo explicado en el apartado “2.2” de esta sección. Todas las muestras fueron previamente dializadas con el mismo tampón: 20 mM Fosfato Sódico pH 6,5; 50 mM NaCl y 1 mM DTT. La mezcla proteica fue incubada durante 1 hora a temperatura ambiente. Finalmente se añadió a cada muestra 10 % de agua

deuterada y 0,16 % de DSS (“2,2-dimethyl-2-silapentane-5-sulfonate sodium salt”) en un volumen final de 600 μ L.

Los espectros bidimensionales fueron realizados por el Dr. Daniel Padró, usando un espectrómetro “Bruker AVANCE 700” a 25 °C de temperatura. Los espectros fueron procesados con el programa *XWINMR* y analizados usando *NMRView*. La interacción entre las moléculas fue analizada en base a la perturbación en los cambios químicos observados en el espectro ^1H - ^{15}N -HSQC de ^{15}N -BRMS1₅₁₋₈₄ al añadir NMI₁₆₋₆₄.

7.3 *Cromatografía de exclusión por tamaño molecular*

- a) Tax1BP1₇₂₅₋₇₈₉ (UBZ1+UBZ2) + Ub
- b) Tax1BP1₇₂₅₋₇₈₉ (UBZ1+UBZ2) + Poli-4Ub
- c) Tax1BP1₇₂₅₋₇₈₉ (UBZ1+UBZ2) + Itch₄₃₅₋₅₁₁ (WW3+WW4)

Las proteínas purificadas de forma individual y concentradas hasta aproximadamente 100 μ M (como se expone en la sección de “*Purificación*”) se incubaron a temperatura ambiente en una proporción equimolar durante 30 minutos. La mezcla proteica se concentró en unidades de filtración de 5 kDa de diámetro de poro (*Millipore*^T), hasta un volumen y concentración detallados en la sección de resultados, y se inyectó en una columna de filtración en gel (del tipo Superdex 75 16/60 ó 26/60). Los tampones finales usados en estas columnas fueron de 20 mM TRIS pH 7,5 y 0,2 M NaCl para el caso del complejo a) y c), y de 20 mM TRIS pH 7,5 y 50 mM NaCl para el caso de b).

Tras la comparación de volúmenes de elución resultantes de la columna de exclusión por tamaño molecular en cada caso; es decir, de cada componente individual por un lado y de la mezcla de ambas proteínas por el otro, se discernieron y aislaron las fracciones correspondientes al volumen de co-elución de ambas moléculas. Analizando el peso molecular correspondiente a cada volumen de elución, obtenido tras interpolarlo en una curva de calibración, se estudió la posibilidad de la formación de complejos. Estas fracciones se concentraron en concentradores de las mismas características que los usados en el paso previo, hasta alcanzar una concentración óptima para ensayos de cristalización (datos cuantitativos expuestos en la sección de *Resultados*).

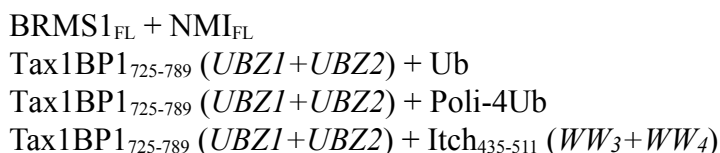
8. *Ensayos de cristalización*

Para abordar el conocimiento de las estructuras de las proteínas objeto de estudio, se empleó la técnica de cristalografía de rayos X, para lo que se hace necesario cristalizar la proteína. En estos ensayos se utilizaron principalmente dos tipos de muestras:

- a) Proteínas individuales, purificadas y concentradas a determinados valores (*Resultados*) tal y como se ha expuesto en los apartados anteriores.
- b) Proteínas en complejo, donde las muestras proteicas procedían de:

- Coelución de ambas proteínas en diferentes pasos de la purificación, tal y como se explica en el apartado 7.3 de esta sección.

Este fue el caso de los complejos:



- Co-cristalización tras incubación post-purificación de ambas proteínas, obtenidas de forma independiente, en un tampón conteniendo 20 mM TRIS pH 7,5 y 50 mM NaCl para BRMS1₅₁₋₈₄ + NMI₁₆₋₆₄ y 20 mM TRIS pH 7,5 y 0,2 M NaCl para Tax1BP1₇₂₅₋₇₈₉ (UBZ1+UBZ2) + Ub. La incubación se mantuvo durante 30 minutos a temperatura ambiente.

8.1 Cristalización inicial

La técnica empleada fue la “difusión de vapor” entre una gota conteniendo la proteína, y un reservorio con la solución de cristalización. Se utilizaron diferentes preparados comerciales en bloques de 96 condiciones (*WIZARD*, *WIZARD cryo* (de Emerald Biosystems^U), *HAMPTON^V Crystal Screen e Index*, *HTSI-II* de Jena Bioscience^W, *NEXTAL Classics-Pegs-Cryos-pHClear-Anions-Cations-JCSG + Suite* (Qiagen^G), y *Natrix*), que se dispensaron en placas de 96 pocillos para ser utilizadas, del tipo *Greiner^Y* (en formato de 1, 2 y 3 gotas por pocillo), *Costar^Z* e *Innovaplate S2-D^{AB}* (en formato de 2 gotas por pocillo).

Se probaron diferentes concentraciones de proteína, proporciones proteína:reservorio y temperaturas (4 °C, 8 °C, 16 °C y 22 °C). Las gotas se pusieron:

- Manualmente, usando una pipeta multicanal. Con un volumen final de la gota entre 1-1,6 μ L. Se probaron diferentes combinaciones: 0,8 μ L proteína + 0,8 μ L reservorio; 0,5 μ L proteína + 0,5 μ L reservorio; 1 μ L proteína + 0,5 μ L reservorio y 0,5 μ L proteína + 1 μ L reservorio.
- Por otro lado, se usó el robot de cristalización “Cartesian MicroSys synQUAD” (*Genomic solutions^X*) para dispensar microgotas de entre 0,2 - 0,4 μ L finales. Se probaron diferentes proporciones entre el volumen de la proteína y el reservorio: 0,1 μ L proteína + 0,1 μ L reservorio; 0,2 μ L proteína + 0,2 μ L reservorio; 0,2 μ L proteína + 0,1 μ L reservorio y 0,1 μ L proteína + 0,2 μ L reservorio.

El volumen del reservorio en todas las placas dispensadas manualmente fue de 50 μL y de 75 μL para las alicuotadas con el robot *Cartesian*

8.2 Optimización de cristales

Una vez encontrada la condición inicial favorable para el desarrollo de cristales, se procedió a su optimización para favorecer el crecimiento y la calidad de los cristales,

intentando conseguir cristales grandes, únicos, con aristas definidas y regulares y de caras uniformes.

Para ello se rastrearon condiciones cercanas variando los parámetros de los componentes iniciales a lo largo de una matriz, creando así un gradiente diagonal. Se usaron placas de 24 pocillos (de *Nextal-Qiagen^G* y *Hampton Research^V*), en forma de “gota colgante”.

En algunos experimentos se usó el robot de cristalización “Tecan freedom evo” de *Tecan^{AC}*, para preparar las soluciones de los reservorios, dispensando 0,5 mL como volumen final en ambos casos.

Para intentar mejorar la calidad de algunos cristales se hizo necesario el uso de diferentes técnicas, como:

– ***Aditivos comerciales***

Se utilizaron por defecto, en todos los casos donde obtuvimos cristales de proteína, diferentes condiciones comerciales de aditivos y detergentes contenidos en bloques (“Additive Screen 1, 2 y 3” y “Detergent Screen 1, 2 y 3”), de la casa comercial *Hampton Research^V*, siguiendo el protocolo establecido.

– ***Ligandos***

Se añadieron a la condición de optimización ciertos componentes, que aunque no estaban en la condición inicial, aparecían como agente común (sales y/o precipitantes) en múltiples condiciones con microcristales, cristales líquidos... Para ayudar a la cristalización y favorecer al plegamiento en α -hélice de BRMS1₅₁₋₈₄ se usó 2,2,2-trifluoroetanol (TFE) a diferentes concentraciones (5, 10, 15 y 20 %), directamente sobre la proteína y/o también sobre el reservorio.

– ***Uso de aceites***

Se intentó ralentizar la difusión de vapor y como consecuencia la aparición y crecimiento de aquellos cristales de desarrollo muy temprano, añadiendo una fina capa de aceites (“Paraffin Oil” y “Silicon Oil”, en español aceite de parafina y de silicona, de la casa comercial *Molecular Dimensions^{AD}*) sobre el reservorio.

Se usaron diferentes proporciones de Parafina/Silicona, de 0 % a 100 % de cada componente con incrementos del 10 % en cada paso. El volumen añadido de la mezcla fue de 100 μ L sobre un reservorio de 0,5 mL conteniendo las condiciones óptimas de cristalización.

– ***Glicerol, isopropanol***

Se añadieron a la condición de cristalización diferentes concentraciones de glicerol e isopropanol para intentar aminorar la nucleación y evitar la formación de cristales frágiles.

– ***Vitrificación-calentamiento-vitrificación*** (en inglés “*annealing*”)

En algunas ocasiones se realizó un rápido calentamiento a temperatura ambiente del cristal ya vitrificado, seguido de nueva vitrificación a 130 K, proceso que en inglés se denomina “annealing”, esto es, “templado”. Se llevó a cabo directamente

sobre el cristal localizado en el goniómetro, por interrupción durante 2 segundos del flujo de nitrógeno gaseoso que lo mantenía vitrificado.

Esta técnica fue probada en cristales con difracción débil o cuya superficie contenía hielo, con el fin de mejorar la difracción a mayor resolución o eliminar los aros generados por el hielo sobre el patrón de difracción (Ref. bibliográfica 3*, pag. 31-42).

– ***Deshidratación del cristal***

Se intentó extraer solvente del cristal siguiendo dos alternativas. La primera basada en una deshidratación manual, exponiendo la gota al aire, sobre aquellos cristales que aguantaron la desecación de la gota que lo contenía. Un segundo método se realizó por sucesivos pasos del cristal o del cubreobjetos conteniendo la gota con el cristal, a través de soluciones de cristalización con concentraciones crecientes del precipitante.

– ***Sembrado de micro- y/o macrocristales*** (en inglés micro- y macro-“*seeding*”)

Con esta técnica se intentó mejorar la calidad de los cristales por deposición de centros de nucleación, generados previamente a partir de cristales de calidad inferior o de material microcristalino, en una nueva gota conteniendo las condiciones de cristalización y proteína en disolución, y que posteriormente se ponen de nuevo a equilibrar contra un reservorio conteniendo las soluciones de cristalización.

Para la realización de la primera técnica (“microseeding”), con ayuda de un pelo de gato soportado a una base, se rozaron cristales ya formados, capturando los centros de nucleación. Seguidamente se pasó el pelo sobre una nueva gota, produciendo de esta forma el micro-sembrado de centros de nucleación. Para la segunda técnica (“macroseeding”) se utilizó un lazo de nylon con el que se cogieron el cristal o los cristales, se “lavaron” en una gota conteniendo la solución de cristalización con el fin de eliminar posibles entrecruzamientos de su superficie y permitir un crecimiento continuado, y se depositaron en la nueva gota.

– ***Diálisis***

Con el fin de probar si la formación de algunos cristales mejoraba con otro método que no fuera el de difusión de vapor, se intentó la cristalización mediante diálisis. Se usaron para ello botones de diálisis de 5 µL (*Hampton Research*^V) con una membrana de 3,5 kDa de diámetro de poro de *Pierce*^S, colocados dentro de placas de 24 pocillos.

9. *Montaje de cristales y vitrificación*

Se seleccionaron cristales únicos, grandes y bien formados. Para manipularlos, se usaron lazos de nylon de diferentes diámetros, desde 0,05 mm hasta 0,5 mm (*Hampton Research*^V). Los lazos y los viales de conservación de los cristales en nitrógeno líquido, se muestran en la misma Figura 8.

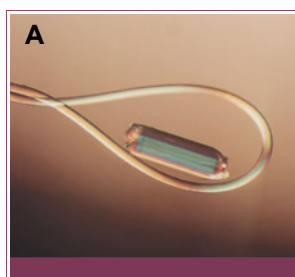


Figura 8. Material usado para el montaje y la conservación de los cristales.

A. Lazo de nylon que soporta al cristal durante su manipulación y difracción.

B. Viales magnéticos de conservación de los cristales en nitrógeno líquido.

El rápido enfriamiento de todos los cristales obtenidos en este proyecto se realizó bajo un flujo de gas de nitrógeno a una temperatura de 130 K, tras colocar la estructura soportando el cristal rápidamente bajo el flujo de gas de nitrógeno.

En algunos casos fue necesario el uso de soluciones crioprotectoras, tales como glicerol, etilénglicol o polietilénglicol para evitar esta formación de hielo cristalino. Los cristales formados en condiciones de cristalización que formaban hielo tras someterse a proceso de vitrificación, fueron bañados previamente en una solución a la que se añadió entre 5 % y 20 % de estos componentes. Aquellos cristales que no soportaron un brusco cambio en la condición, fueron sometidos a baños sucesivos conteniendo concentraciones crecientes del elemento crioprotector hasta alcanzar aquella óptima para la vitrificación.

En algunos casos se hizo necesario el uso de aceites de alta densidad (como Paratone-N) para desolvatar el cristal.

10. Resolución de estructuras

10.1 Recolección de datos de difracción

Tras el proceso de vitrificación, los cristales fueron irradiados por un haz de alta intensidad de rayos X. Se usaron dos tipos de fuentes: Una fuente de rayos X de ánodo rotatorio *Bruker F530* localizada en el Centro Nacional de Investigaciones Oncológicas (CNIO), con un detector del tipo *mar345 (MarResearch)* y otra fuente de radiación de sincrotrón localizada en Grenoble (ESRF, “European Synchrotron Radiation Facility”), donde se usaron las líneas ID14-1, ID23-1 y BM16 (Tabla V), y los detectores “*ADSC Quantum Q210*” y “*ADSC Quantum Q315r*” y “*ADSC Q210r Reversephi*” respectivamente.

Línea	Tipo de energía	Rango de energía	Tamaño del haz (μm)	Límites de resolución (Å)	Tiempo de exposición típico	Velocidad del ángulo φ recomendado	Número de pases
ID14-1	Fija	0,934 Å (13,274 keV)	100	0,95	1 - 5 seg.	0,3 seg./grado	3
ID23-1	Sintonizable	0,62 a 2,5 Å (20 a 5,0 keV)	50	0,64	0,1 - 1 seg.	0,1 seg./grado	1
BM16	Sintonizable	0,73 a 2,06 Å (17 a 6 keV)	100 - 200	0,7	1 - 20 seg.	1 grado/seg	-

Tabla V. Líneas del sincrotrón de Grenoble donde se llevó a cabo la difracción y recolección de los datos. Se muestran las características generales (<http://www.esrf.eu>) de las líneas usadas, así como el rango óptimo de parámetros aconsejados para la realización de experimentos (tres últimas columnas).

Las imágenes conteniendo los datos originados de la exposición de un cristal a difracción por rayos X se recogieron usando el programa informático de recolección de datos de *MarResearch*, en la fuente de rayos X del CNIO, mientras que en el sincrotrón se utilizó el programa *MXCuBE* (“MX Customised Beamline Environment”). Para optimizar la recolección de juego de datos de modo que estos fuesen completamente informativos; es decir, que contuvieran para una resolución determinada el 100 % de las reflexiones esperadas y con una relación señal/ruido adecuada, se usó una herramienta informática “*Strategy*”, dentro del paquete informático “*Mosflm*”. Con este programa se trazaron planes estratégicos de recolección para obtener una elevada completitud y multiplicidad ajustada al tipo de juego de datos, lo que conlleva usar unos parámetros determinados para cada estructura (expuestos para cada caso en el apartado de “*Resultados*”).

10.2 Reducción de datos

La simplificación de los datos se realizó usando el programa informático *Mosflm* para el indexado, y con ello determinar los parámetros de la celda unidad del cristal, y el integrado de las reflexiones, y *Scala* del paquete informático *CCP4* (CCP4, 1994), para el escalado y combinado. La simetría del cristal fue estimada inicialmente en el proceso de indexado, y posteriormente con “*pointless*” y con la ayuda de la “*Self Rotation Function*”, implementada en *Molrep* (CCP4).

Las estimaciones sobre la cantidad de solvente en el cristal, así como del número de moléculas dentro de la unidad asimétrica, se realizaron usando la herramienta *matthews_coef* del paquete *CCP4i* (CCP4, 1994); esto es, calculando el coeficiente de Matthews (Kantardjieff y Rupp, 2003; Matthews, 1968).

10.3 Resolución de fases experimentales

Para resolver las fases de las estructuras resueltas en este proyecto seguimos dos aproximaciones: *SIRAS*, tras obtener un cristal “derivado”o con señal anómala de un átomo pesado, para el caso de BRMS1₅₁₋₈₄, y *SAD*, utilizando la señal anómala procedente

de átomos de zinc presentes en el cristal nativo, para el caso de UBZ1+UBZ2 de Tax1BP1.

10.3.1 Obtención de cristal derivado de átomo pesado

La selección del derivado de átomo pesado óptimo se realizó a partir de un “kit” comercial de *Hampton Research*^V formado, en total, por 37 derivados de átomos pesados, de Platino, Mercurio y Oro.

En todos los cristales obtenidos dentro del proyecto donde se pretendía incorporar un metal pesado se realizaron tres rondas de selección para elegir aquellos mas viables para la interacción con la molécula y con ello reducir el número de átomos a probar sobre los cristales:

✓ *Selección del rango de pH estable*

La casa comercial proporcionó junto al producto una lista detallada del rango de pH en el que cada derivado es estable. Se descartaron los derivados de metales pesados cuyo rango de pH no abarcase al de la condición de cristalización donde se formaban los cristales que se iban a impregnar.

A su vez, también se descartaron los derivados que no fuesen solubles en dicha solución de cristalización.

✓ *Afinidad por diferentes residuos*

Examinando la secuencia aminoacídica de la proteína, se estudió la presencia de ciertos residuos por los que los metales pesados tienen afinidad (por ejemplo, el mercurio y el oro son afines a residuos de cisteína libres y a histidinas, mientras que el platino lo es para residuos de metionina, histidinas, cisteínas y residuos con carga positiva), y se descartaron los derivados que podían tener menos afinidad.

✓ *Diferencias en la migración en un gel nativo*

Para descartar aquellos derivados, de los seleccionados hasta aquí, que no interaccionasen con la proteína, se incubó durante 30 minutos la proteína en la solución de cristalización (la formación de cristales ocurría de 3 a 10 días mas tarde, según el caso) con concentraciones en exceso de cada derivado de átomo pesado, individualmente. A continuación se añadió 4 % de glicerol a estas muestras y se cargaron en un gel nativo al 17 % de poliacrilamida. El tampón de carrera usado contenía 5 mM TRIS y 0,3 % glicina. La carrera del gel se mantuvo a un voltaje constante de 150 V, durante 4 horas a una temperatura de 4 °C.

Tras teñir las bandas de proteína con azul de Coomassie, se comparó el patrón de migración de la proteína con cada átomo, de forma que la mezcla que no había entrado en el carril, se consideró precipitada o desnaturalizada, las que mantenían el mismo patrón que la proteína control (sin derivado) se descartaron,

mientras que los derivados que cambiaron el patrón de migración de la proteína fueron considerados para incorporarlos al cristal (ver “*Resultados*”).

Una vez seleccionado el derivado de metal pesado, se realizó la impregnación o empapado del cristal en una solución conteniendo el átomo pesado (en inglés “soaking”). Se probaron diferentes concentraciones de átomos al igual que distintos tiempos de incubación.

Para ello, o bien se complementó la gota conteniendo el cristal (generalmente de 2 μ L de volumen) con diferentes volúmenes del derivado pesado disuelto en la condición de cristalización, quedando a concentraciones finales de entre 0,1 mM y 10 mM, o bien se traspasó el cristal a una gota fresca con la condición de cristalización, conteniendo el átomo pesado.

Estas incubaciones se mantuvieron también durante diferentes intervalos de tiempo: 5, 15 y 30 minutos y 1, 2 y 5 horas y toda la noche.

Estos parámetros (concentración y tiempo) fueron combinados aleatoriamente, hasta obtener los mayores valores a los cuales el cristal no sufría daños tanto externos como internos (ver “*Resultados*”).

Tras la impregnación, el cristal con el derivado de átomo pesado fue lavado 3 veces en gotas limpias conteniendo la solución de cristalización para eliminar la presencia residual de metal pesado en la superficie del cristal y posteriormente se vitrificaron y montaron en la fuente de rayos X tal y como se expuso en el apartado número 9.

10.3.2 Métodos de resolución de estructura usando señal anómala

a) SIRAS (“Single Isomorphous Replacement with Anomalous Scattering”)

Para resolver las fases se usaron las diferencias anómala e isomorfa originadas por un único derivado de átomo pesado medido a la longitud de onda específica donde absorbe rayos X.

Tras la reducción de datos procedentes de la difracción de un cristal nativo y otro derivado, se utilizaron algunas herramientas de la interfaz informática *CCP4i* (CCP4, 1994) para estudiar la existencia de señal anómala. Estos juegos de datos se manipularon y combinaron usando los programas *Sftools* y *CAD* y se escalaron con *Scaleit* para calcular un “mapa de Patterson” con las de diferencias isomorfas (“*Patterson, run FFT*”).

Posteriormente los juegos de datos se trataron por separado para generarlos en formato “Scalepack”, legible por el programa informático *SHELX* (Sheldrick, 2008) (bajo la interfaz *hkl2map 0.2*) con el que se resolvieron las fases. Primero los datos experimentales se trataron y prepararon para las etapas posteriores, usando *SHELXC*. Con *SHELXD* se localizaron las posiciones de los átomos pesados mediante métodos directos en el espacio real y recíproco con posiciones originadas a partir del “mapa de Patterson” generado con la señal anómala. Finalmente, se utilizó el programa *SHELXE* para calcular las fases y realizar una primera modificación de la densidad electrónica, con objeto de mejorar los mapas de densidad iniciales.

b) SAD (“Single-wavelength Anomalous Diffraction”)

Para resolver las fases, en este caso se utilizó exclusivamente la señal anómala recogida en un juego de datos de un cristal derivado, a una única longitud de onda. Los datos se procesaron con el programa *autoSHARP* (Vonrhein *et al.*, 2007).

10.4 Construcción del modelo

Para la construcción de un modelo inicial de la estructura, se utilizaron los programas “ARP/wARP” (Perrakis *et al.*, 1999) y “Buccaneer” (Cowtan, 2006) del paquete CCP4 (CCP4, 1994) y “Resolve de PHENIX (AutoBuild wizard)” (Adams *et al.*, 2002; Terwilliger, 2002; Terwilliger *et al.*, 2008).

Cuando fue preciso combinar diferentes modelos para originar uno nuevo y editar alguno de ellos, se utilizaron las herramientas *Coot* (Emsley y Cowtan, 2004), *pdbsset* del paquete CCP4 (CCP4, 1994), *Moleman* y editores de texto con extensiones para editar ficheros de coordenadas, *pdb*s, tales como *emacs*.

10.5 Afinado de la estructura

Para ajustar el modelo a los datos experimentales, se realizaron afinamientos isotrópicos y anisotrópicos.

Para llevar a cabo el afinamiento de la molécula se utilizó *REFMAC5* (Murshudov *et al.*, 1997). Dentro de este programa se realizaron diferentes tipos de afinamientos con distintos parámetros según la estructura: *Afinamiento de cuerpo rígido* (en inglés “Rigid Body”), *simetría no cristalográfica* (NCS) y afinamiento usando *TLS* (Winn *et al.*, 2001).

Los “grupos de TLS” fueron seleccionados usando la herramienta informática *TLSMD* (Painter y Merritt, 2006).

Para el afinamiento de cada modelo se intercalaron afinamientos de las moléculas en el espacio real utilizando *Coot* (Emsley y Cowtan, 2004), combinados con afinamiento en el espacio recíproco con *Refmac5* (Murshudov *et al.*, 1997).

Finalmente se colocaron moléculas de agua con la herramienta destinada para ello facilitada por *Coot* (Emsley y Cowtan, 2004), y de forma manual.

10.6 Validación de la estructura

Se chequearon parámetros como la correlación en el ajuste de densidad, errores en las coordenadas, los factores *R* y *R*_{free} o los factores *B* del modelo mediante herramientas como *Sfcheck*, *Procheck* y *Baverage*, dentro del paquete CCP4, así como las desviaciones

cuadráticas medias (en inglés, “*r.m.s.d.*” por “root mean square deviation”) sobre los ángulos y la longitud de los enlaces (parámetros dados por *Refmac5*).

Los ángulos “phi; ϕ ” y “psi; ψ ” permitidos entre los enlaces diedros de la cadena principal C_{i-1} -Ni- $C\alpha_i$ - C_i y N_i - $C\alpha_i$ - C_i - N_{i+1} respectivamente fueron chequeados mediante un diagrama de “*Ramachandran*” (Kleywegt y Jones, 1996; Lovell *et al.*, 2003), facilitado por el programa *Coot* (Emsley y Cowtan, 2004).

Con otras utilidades del programa *Coot* (Emsley y Cowtan, 2004) comprendidas dentro de la herramienta “*Validate*”, se corrigió la quiralidad de algunos de los carbonos de los residuos, se analizaron la geometría, el ajuste en la explicación de la densidad electrónica y los factores de temperatura de los rotámeros y se comprobó la validez de las aguas modeladas en la estructura.

Finalmente se analizaron los contactos entre átomos y la geometría con *MolProbity* (Davis *et al.*, 2007).

Para el estudio y caracterización de las estructuras resueltas se utilizaron programas como *Twister* (Strelkov y Burkhard, 2002), para el análisis de motivos “coiled-coils”, *Monster* (Salerno *et al.*, 2004) y *Contact* de *CCP4*, para analizar los contactos entre residuos y *AreaIMol* (*CCP4*), para estudiar las superficies de contacto intermoleculares.

Resultados

La cristalografía de macromoléculas es una técnica compleja que precisa de unos niveles de pureza tanto química como estructural muy exigentes para intentar el abordaje de las series de cristalización. Por ello, se procedió a una selección cuidadosa de distintos fragmentos de las proteínas objeto de estudio, así como de las proteínas completas, que se expondrán en los respectivos capítulos. Se utilizaron los genes humanos.

Para la selección y diseño de las proteínas objeto de interés, se realizó una búsqueda de la información existente de cada una de ellas en diversas bases de datos: *ExPASy*³, *UniProtKB/Swiss-Prot*⁴, *PROSITE*², *BMERC*⁵, *Bioinformatic Harvester (EMBL)*⁶, *NCBI*⁷, *GeneCards*⁸, *Human Protein Reference Database (hprd)*⁹ o *PDB*¹⁰. Con estas herramientas informáticas se estudiaron las características biológicas, funcionales, físico-químicas y estructurales conocidas de dichas proteínas humanas, y que se expusieron en la *Introducción*.

Las dos proteínas en torno a las cuales se desarrollaron los estudios son BRMS1 y Tax1BP1, concretamente se trabajó con las isoformas expuestas en *negrita* en la Tabla VI; esto es, la *isoforma 1* para BRMS1, de 246 residuos de longitud y la *isoforma a* para Tax1BP1, de 789 aminoácidos.

Proteína	Isoformas	Código (gen codificante)	Código (Proteína)	Peso Molecular	Número residuos
BRMS1	- Isoforma 1: 246 aa - Isoforma 2: 290 aa - Isoforma 3: 152 aa	- NM_015399 - NM_001024957 - NM_001024958	- NP_056214 - NP_001020128 - NP_001020129	28.461 Da (Isoforma 1)	246 aa (Isoforma 1)
Tax1BP1	- Isoforma a: 789 aa - Isoforma b: 747 aa	- NM_006024 - NM_001079864	- NP_006015 - NP_001073333	86.257 Da (Isoforma 2)	747 aa (Isoforma 2)

Tabla VI. Isoformas descritas para las proteínas BRMS1 y Tax1BP1. Sus respectivos códigos de acceso tanto para la secuencia de ADN codificante como para la proteína, respecto a la base de datos *NCBI*⁷, se exponen en las columnas centrales. Se muestran también el peso molecular y el tamaño molecular solamente de la isoforma usada en este proyecto. En *negrita* se resalta la información correspondiente a la isoforma de cada proteína, objeto de estudio. Abreviatura de aminoácidos: aa.

Capítulo 1

BRMS1.

Resumen

En este capítulo se expone el clonaje, expresión y cristalización de las construcciones diseñadas de la proteína supresora de metástasis *BRMS1*. Así mismo se muestra la caracterización cristalográfica de su primer motivo “*coiled-coil*”, que media interacción con *NMI*, una proteína de unión a factores de transcripción como *N-myc* y *c-Myc*, y con ello cobra un papel fundamental en la función biológica de *BRMS1*.

Los ensayos de caracterización de la interacción de *BRMS1* con el co-factor de transcripción *NMI* realizados en este proyecto, sugieren un posible mecanismo de supresión de metástasis.

1. *Diseño de las construcciones objeto de estudio*

Con el fin de abordar el diseño de las construcciones de BRMS1, se realizó en primer lugar un análisis de la predicción de estructura secundaria de la secuencia aminoacídica humana, utilizando el programa *PSIPRED*¹¹, lo que mostró que la estructura se compone mayoritariamente de α -hélices, además de regiones no estructuradas (ver Figura 9). El alineamiento entre diversas especies muestra una elevada homología en la mayor parte de la secuencia aminoacídica, demostrando una alta conservación evolutiva de BRMS1, tan solo diferenciándose en residuos puntuales (Figura 9). El fragmento de aminoácidos 185-200 es el que muestra mas divergencia entre estas especies. El alineamiento se realizó con *MultAlin*¹³, tras una búsqueda de ortólogos para cada gen (*GeneCard*⁸ y *HomoloGene* (NCBI)¹²).

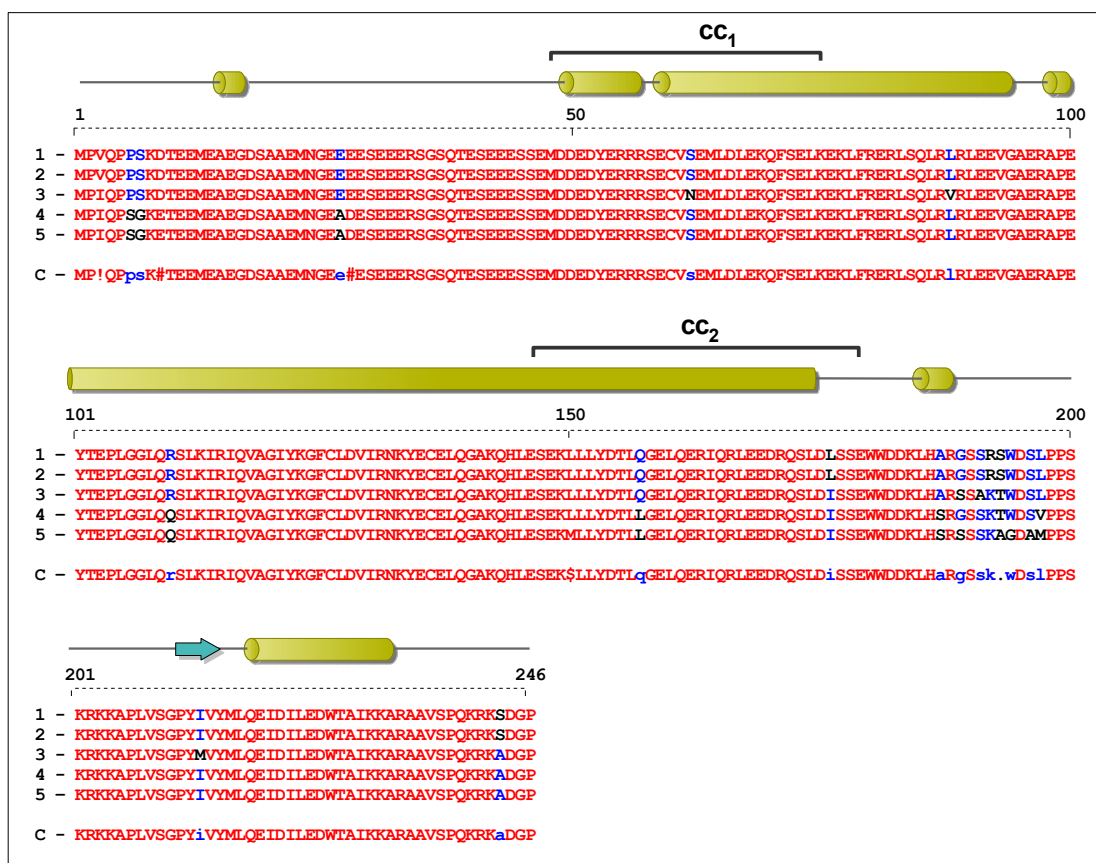


Figura 9. Alineamiento proporcionado por el programa informático *MultAlin*¹³ basado en la secuencia aminoacídica de la proteína BRMS1 entre diferentes especies: **1** - *Homo sapiens* (Q9HCU9), **2** - *Pan troglodytes* (XP_001171253 - Isoforma 1), **3** - *Canis lupus familiaris* (XP_854736), **4** - *Rattus norvegicus* (NP_001009605), **5** - *Mus musculus* (NP_598916), **C** - Secuencia consenso.

Se muestran en rojo los aminoácidos mas conservados, en azul residuos de nivel medio de homología y en negro aquellos que se encuentran menos conservados. Los símbolos !, ., # hacen mención a la tabla de *Blosom62*, significan que aunque no se trata del mismo aminoácido, se encuadran dentro del grupo con las mismas propiedades físicas, !=Valina e Isoleucina y \$=Metionina y Leucina (hidrofóbicos, apolares), #=Aspártico y Glutámico (polares, de carga negativa) y .=aminoácido no conservado.

Sobre el alineamiento se muestra la predicción de estructura secundaria aminoacídica de BRMS1 facilitada por el programa *PSIPRED*¹¹. Con forma de cilindros se muestran las estructuras en α -hélice, mientras que las flechas corresponden a láminas β . El resto de las regiones (representadas con una línea) se corresponden con regiones sin estructura definida. Con los programas *Coiled-coil prediction*¹⁴, *Multicoil*¹⁵ o *pbil*¹⁶ se predijeron dos motivos “coiled-coil” (CC₁ y CC₂), comprendidos entre los aminoácidos 51-78 y 147-179)

Utilizando otros programas de predicción de motivos “coiled-coil” (*Coiled-coil prediction*¹⁴, *Multicoil*¹⁵ o *pbil*¹⁶) comprobamos, como ya se describió previamente en la literatura (Seraj *et al.*, 2000b), la existencia de dos de estos motivos, comprendidos entre los aminoácidos 51-78 y 147-179, que coinciden con predicción de estructuras helicoidales como se puede observar en la Figura 9.

Con toda esta información, nos centramos en el diseño de construcciones que contuvieran regiones estructuradas completas, tanto helicoidales como de β -lámina. Por otro lado también nos centramos en el clonaje y expresión de motivos “coiled-coil” por dos razones: La primera está basada en resultados obtenidos previamente en el laboratorio por el Dr. José Rivera quien, mediante ensayos de doble híbrido, comprobó la importancia del primer motivo “coiled-coil” de BRMS1 como mínima región de interacción con diversas proteínas de gran interés biológico (artículo en revisión). La segunda razón es la existencia de pocos motivos estructurales reconocibles con importancia biológica, como se muestra en la Figura 3, a lo que se sumó la importancia del papel que desempeñan estos motivos en interacciones entre proteínas y en funciones biológicas de gran relevancia (Mason y Arndt, 2004).

Una de las proteínas rescatadas en los ensayos de doble híbrido mencionados anteriormente, es NMI (“N-Myc and STAT interactor”), un co-factor de la transcripción que interacciona con diversos transductores y activadores de este proceso, como N-Myc, c-Myc o proteínas de la familia STAT modulando importantes vías de señalización (Bannasch *et al.*, 1999; Zhu *et al.*, 1999).

Como se observa en la Figura 10, NMI consta de un fragmento amino-terminal con elevada homología con la proteína de 35 kDa inducida por interferón, que contiene una región helicoidal rica en leucinas; aminoácidos 29-104 (dominio *IFP35*), dos dominios de aproximadamente 90 residuos (*NID*) que median la homodimerización NMI-NMI y localización subcelular, y un motivo predicho en “coiled-coil” entre los residuos 21 a 64 (Chen *et al.*, 2000).

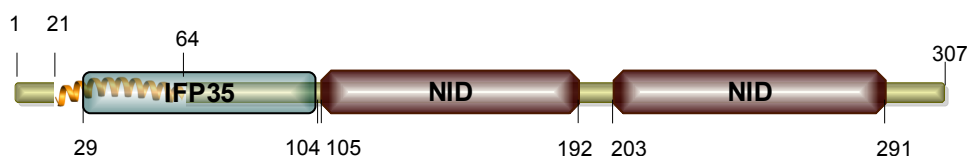


Figura 10. Representación esquemática de la predicción de dominios de la proteína NMI humana. La predicción, facilitada por los programas *Multicoil*¹⁵ y *Pfam*¹⁹, muestra la existencia de una región similar al amino terminal de la proteína de 35 kDa inducida por interferón (*IFP35*), de dos dominios que median interacción NMI-NMI y de localización subcelular (*NID*) y de un motivo “coiled-coil”. Se muestran en números todos los aminoácidos limitantes de estas regiones.

En concreto, los dominios mínimos de NMI de interacción con BRMS1 rescatados en el ensayo de doble híbrido fueron, por un lado el comprendido entre los residuos 1 al 104, que abarca los motivos “coiled-coil” e *IFP35* y, por otro lado, los dominios *NID*. La presencia en NMI de una estructura predicha como “coiled-coil” en el primer fragmento rescatado, nos llevó a pensar en la posibilidad de una interacción mediada únicamente por dichos motivos, ya que está descrito desempeñan un papel muy importante en diversas funciones biológicas y en interacción proteína-proteína (Lupas, 1996). Por este motivo, centramos nuestro interés en el estudio estructural del primer “coiled-coil” de BRMS1 individual y de su interacción con la región “coiled-coil” de NMI. Por otro lado, parte de este motivo predicho está contenido en el dominio *IFP35*, por lo que también nos centramos en el estudio del fragmento *IFP35*. Para dichos análisis, diseñamos numerosas construcciones, ayudándonos de la predicción de estructura secundaria con el fin de intentar evitar dominios sin estructurar.

En la Tabla VII se muestran las construcciones diseñadas para BRMS1 y NMI. Además de centrarnos en las secuencias completas de ambas proteínas, diseñamos diversos clonajes para BRMS1 que contenían, o bien alguno de los dos motivos “coiled-coil”, o los dos. Algunas de las construcciones con el primer motivo (48-78, 48-84, 51-78 y 51-84) se diseñaron mostrando tan sólo 5 residuos de diferencia entre sí en alguno de sus extremos, con el fin de incrementar las probabilidades de solubilización permitiendo la formación de diferentes estructuras en los extremos, para abarcar un mayor margen de error, ya que no debemos olvidar que se trata de una predicción. El resto de las construcciones fueron diseñadas intentando respetar los dominios estructurados en α -hélice, mostrados en la Figura 9.

En lo que respecta a NMI, y al igual que BRMS1, todas contenían el motivo predicho como “coiled-coil”. Abordando el mismo método que anteriormente descrito, se diseñaron las construcciones 16-64, 16-69, 21-64 y 21-69. Otros fragmentos (20-52, 24-59 y 30-63, mas cortos que el motivo predicho) fueron ideados para intentar asemejar la longitud del primer “coiled-coil” de BRMS1. El resto de construcciones se diseñaron abarcando también el motivo *IFP35*, alargando la secuencia en algunos casos (fragmentos 1-197 y 1-111) con el fin de introducir regiones estructuradas (predicción de estructura secundaria no mostrada).

2. Resultados generales de las construcciones clonadas

Los fragmentos diseñados fueron clonados en diversos vectores procariotas (pET o pGEX), para su expresión. Tras la optimización de los diferentes protocolos de expresión, como se señaló en el apartado de *Materiales y Métodos*, se obtuvieron adecuados niveles de expresión proteica, que nos permitieron el abordaje de estudios biofísicos y cristalográficos, donde se requiere de una cantidad adecuada y un elevado grado de pureza.

Se obtuvieron diferentes comportamientos en cuanto a la solubilidad o estado de agregación de cada construcción, en cada una de dichas etapas, por ello no todas las construcciones alcanzaron la última fase de cristalización. Estos resultados se resumen en la Tabla VII.

	A	B Construcción	C Purificación	D Estudio
BRMS1	cc1: 51 _E -78 _L	aa (48 _M - 78 _L)	Soluble	
		aa (48 _M - 84 _S)	No expresión	
		aa (51 _E - 78 _L)	Soluble	
		aa (51 _E - 84 _S)	Soluble*	Cristalizado / Estructura
		aa (51 _E - 98 _A)	Soluble*	Cristalizado
	cc2: 147 _S -179 _E	aa (1 _M - 88 _L)	Soluble	
		aa (63 _E - 188 _R)	Insoluble	
	cc1: 51 _E -78 _L cc2: 147 _S -179 _E	aa (63 _E - 246 _P)	Insoluble	
		FL: aa (1 _M - 246 _P)	Agregado*	Cristalizado en presencia de NMIh FL
		aa (1 _M - 187 _A)	Agregado	
		aa (13 _M - 187 _A)	Agregado	
		aa (13 _M - 241 _R)	Insoluble	
		aa (52 _D - 187 _A)	Insoluble	
		aa (52 _D - 241 _R)	Insoluble	
		aa (48 _M - 188 _R)	Insoluble	
		aa (48 _M - 246 _P)	Insoluble	
NMI	cc1: 21 _I -64 _I	aa (16 _S - 64 _I)	Soluble*	Ensayos de interacción con BRMS1 ₅₁₋₈₄
		aa (16 _S - 69 _M)	Insoluble	
		aa (21 _I - 64 _I)	No expresión	
		aa (21 _I - 69 _M)	No expresión	
		aa (20 _F - 52 _O)	Soluble	Cristalizado en presencia de BRMS1 ₅₁₋₈₄
		aa (24 _E - 59 _O)	Soluble	
		aa (30 _I - 63 _D)	Soluble	
	cc1: 21 _I -64 _I IFP35: 29-104 _Q	FL: aa (1 _M - 307 _E)	Agregado*	Cristalizado en presencia de BRMS1 FL
		IFP35 domain (aa 1 _M - 104 _O)	Soluble	Cristalizado en presencia de BRMS1 ₅₁₋₈₄
		aa (1 _M - 197 _D)	Insoluble	
		aa (1 _M - 111 _K)	Soluble	

Tabla VII. Resumen de los resultados obtenidos para las construcciones de las proteínas humanas BRMS1 y NMI. En la columna **A** se muestra la predicción de motivos “coiled-coils” y dominio IFP35. **B)** Se exponen los aminoácidos limitantes en los extremos amino y carboxilo terminales de cada fragmento y en los subíndices, el residuo correspondiente en código de una letra. En la siguiente columna, **(C)** se resume el resultado obtenido tras la purificación: No expresión, Insoluble o Soluble tras la lisis bacteriana o Agregado en una columna de exclusión por tamaño molecular. Las construcciones cuyos resultados se detallan en este capítulo, están representadas con un asterisco y sobre fondo de color. **D)** Se muestra la finalidad y el logro obtenido con dichas construcciones, tanto si fueron cristalizadas (fondo amarillo) o se resolvió su estructura (fondo verde), o fueron destinadas a caracterización de interacción proteína-proteína (fondo naranja).

FL: secuencia completa de la proteína. **aa:** aminoácidos.

Los diferentes resultados obtenidos, reflejados en la Tabla VII, en las diferentes etapas de la expresión, purificación y cristalización, se resumen a continuación:

Algunas construcciones mostraron niveles de expresión no factibles para la obtención de proteína destinada a cristalización. De aquellas proteínas que sí se obtuvieron elevados niveles de sobre-expresión se procedió a la ultracentrifugación del lisado bacteriano, tras la cual, algunas resultaron insolubles o localizadas en la fracción sedimentada, tras intentar evitarlo mediante la variación de agentes químicos que favorecen la solubilidad. Los fragmentos que fueron solubles, se purificaron siguiendo los pasos expuestos en la sección de *Materiales y Métodos*, donde algunas de ellas resultaron agregadas, tras discernir su tamaño molecular en una columna de exclusión o filtración en gel. La composición de los tampones de las proteínas que resultaron agregadas fueron en ambos casos 20 mM BIS-TRIS pH (5,8-7,0) ó 20 mM TRIS pH (7,0-9,0); (0,1-0,5) M NaCl; +/- 2 mM DTT y +/- 4 % glicerol, así como en 20 mM Fosfato Sódico pH 7,4; (0,1-0,5) M NaCl; +/- 2 mM DTT y +/- 4 % glicerol.

Finalmente, tras obtener las proteínas en elevado grado de pureza, sólo algunas llegaron a cristalizar con éxito. Concretamente se consiguieron cristales del primer motivo “coiled-coil”, perteneciente a los fragmentos 51-84 y 51-98 de la proteína BRMS1, así como de la mezcla proteica formada por las secuencias completas de BRMS1 y NMI. También se obtuvieron cristales del motivo “coiled-coil” (20-52) y el IFP35 (1-104) de NMI junto con el motivo “coiled-coil” de BRMS1. Estos últimos cristales no pudieron ser usados para estudios cristalográficos, debido a la calidad de los mismos.

Para simplificar los resultados obtenidos, se exponen solamente los detalles de expresión, purificación y cristalización de aquellas construcciones que nos permitieron obtener datos discutibles e importantes para el desarrollo de este estudio. Éstas aparecen señaladas con un asterisco (*) en la Tabla VII, donde se describen los resultados obtenidos, bien porque se obtuvo la estructura (es el caso de BRMS1 aa51-84) o bien porque se obtuvieron cristales o fueron destinadas a la formación de complejos y estudios de caracterización bioquímica (BRMS1 FL (secuencia completa), BRMS1 aa51-98, NMI FL (secuencia completa) y NMI aa16-64).

Así, los siguientes resultados de estas construcciones particulares se hacen extensivos a los demás fragmentos solubles que llegaron a la fase de cristalización y que fueron clonados con la misma etiqueta de fusión (bien sea histidinas o GST).

3. ***Expresión, purificación, caracterización y cristalización del primer motivo “coiled-coil” de BRMS1 (BRMS1₅₁₋₈₄) (Spínola-Amilibia et al., 2008)***

3.1 ***Expresión***

El ADN codificante para la región comprendida entre los aminoácidos E51 y S84 fue clonado, tras amplificación mediante PCR y posterior digestión según se describe en *Materiales y Métodos*, en el vector pET28-SMT3, para ser expresada en un sistema

heterólogo procariota, originando una proteína de fusión unida a 6 histidinas y a la proteína SMT3 en su extremo amino-terminal (Figura 11). Tras incubación con la proteasa ULP, se liberó un fragmento mayoritario de 122 residuos, correspondiente a His-SMT3, obteniendo de esta forma el polipéptido que denominaremos BRMS1₅₁₋₈₄. Esta secuencia polipeptídica de 34 residuos, que comienza con una serina adicional originada tras el corte con la proteasa y numerada como 0, se utilizó para ensayos de cristalización.

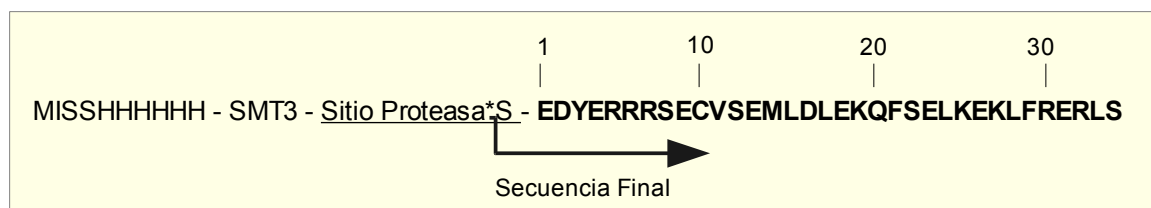


Figura 11. Secuencia originada del clonaje del fragmento comprendiendo los aminoácidos 51 al 84 de BRMS1 antes y después de su proteólisis. Se muestra la secuencia aminoacídica originada tras su expresión, que consta de una cola de 6xHis-SMT3 fusionada en el extremo amino terminal de la proteína, que fue eliminada tras la proteólisis con “ULP” (su sitio de reconocimiento se presenta subrayado), originándose la secuencia final mostrada. Con un asterisco se muestra el sitio de corte de la proteasa. La secuencia en “negrita” es la delimitada por los aminoácidos Glutámico 51 (residuo 1 en la figura) y Serina 84 (residuo 34 en la figura) de BRMS1.

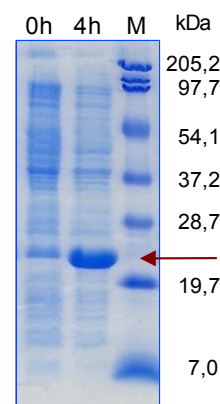
Tras la optimización de las condiciones de expresión (mostradas en la sección de *Materiales y Métodos*), se obtuvieron elevados niveles de sobre-expresión de la proteína His-SMT3-BRMS1₅₁₋₈₄ usando *Rosetta(DE3)* como cepa de *E.coli* de expresión bajo los siguientes parámetros: crecimiento del cultivo en medio *2xYT*, a 37 °C e induciendo la expresión con 0,5 mM de IPTG. El cultivo se recolectó mediante centrifugación transcurridas 4 horas de la adición del agente inductor. En la Figura 12 se compara visualmente el nivel de proteína sobre-expresada, mediante electroforesis en gel de poliacrilamida (SDS-PAGE) de las muestras recogidas antes y después de la adición de IPTG al medio de crecimiento.

Figura 12. Expresión de His-SMT3-BRMS1₅₁₋₈₄ en cepa de *E.coli*, analizada mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 17 % TRIS-HCl.

El carril marcado como “0h” se corresponde con el total de proteínas antes de inducir su sobre-expresión de His-SMT3-BRMS1₅₁₋₈₄ mediante adición de 0,5 mM de IPTG. Transcurridas 4 horas, a 37 °C de temperatura, se recogió el cultivo celular expresando la proteína de interés (carril “4h”), para ser purificada.

Los marcadores (M) de alto peso molecular usados (*BioRad*^E), se muestran a la derecha en unidades de kDa.

La banda correspondiente a la proteína His-SMT3-BRMS1₅₁₋₈₄ sobre-expresada, de un tamaño aproximado de 18,2 kDa, se indica con una flecha roja.



3.2 Purificación

La masa celular obtenida tras la expresión fue lisada mediante sonicación en un tampón de lisis 50 mM TRIS pH 7,5; 0,5 M NaCl, complementado con 0,5 % (v/v) Tween-20 para favorecer la solubilización de la proteína sobre-expresada. Este protocolo nos permitió conseguir prácticamente la totalidad de His-SMT3-BRMS1₅₁₋₈₄ soluble, tras un proceso de ultracentrifugación, como se muestra en la Figura 13. Dicha fracción fue sometida a diferentes etapas de purificación, detalladas a continuación.

3.2.1 Cromatografía de afinidad

El sobrenadante obtenido tras la lisis y ultracentrifugación (“S” en la Figura 13) se filtró y a continuación se purificó mediante cromatografía de afinidad por histidinas, utilizando una columna quelante de histidinas, cargada con sulfato de níquel. Con el fin de obtener el mayor grado de pureza, la proteína unida a la columna se lavó con un tampón 50 mM TRIS pH 7,5; 0,15 M NaCl, y posteriormente se aplicó un gradiente lineal de Imidazol hasta 0,5 M, para que las proteínas eluyeran (Tabla III-C). Este protocolo nos permitió obtener dos máximos de absorbancia, de los cuales el segundo eluido con 168 mM de Imidazol (33 % del tampón B), observamos que contenía la mayor parte de la proteína de interés con un baja proporción de proteínas contaminantes (Figura 13), tras análisis mediante electroforesis SDS-PAGE, por lo que no fue necesaria la utilización de ninguna otra columna, como paso intermedio, antes de tratar la muestra con proteasa.

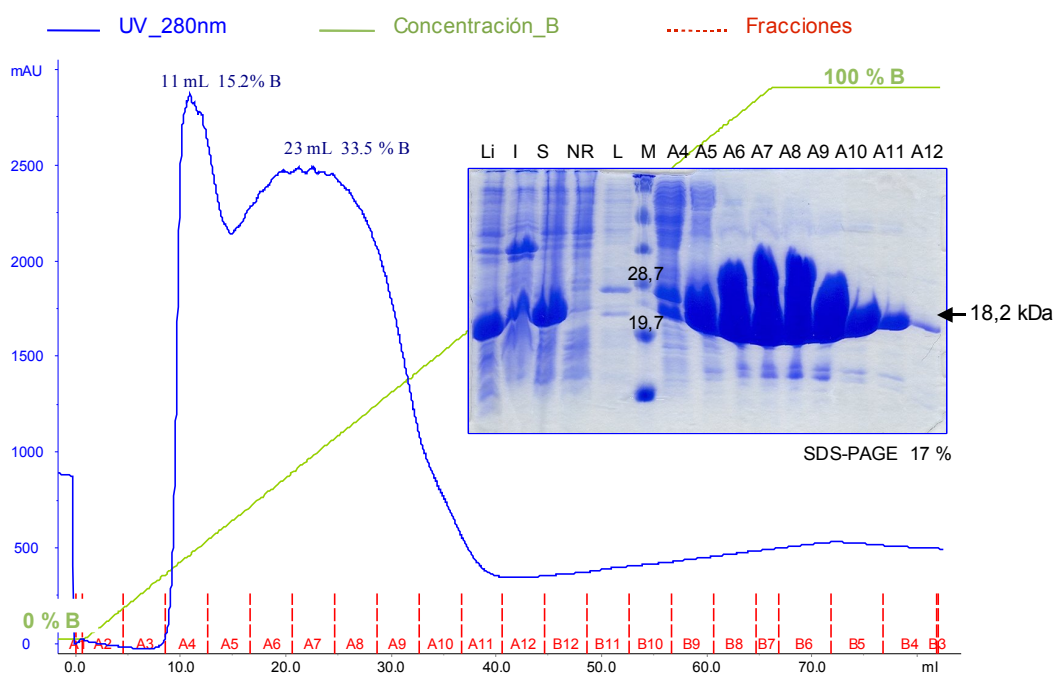


Figura 13. Cromatograma y muestras resultantes de la columna de afinidad a histidinas en una purificación de His-SMT3-BRMS1₅₁₋₈₄. El lisado celular obtenido tras la sobre-expresión de la proteína se cargó en una columna de afinidad precargada con níquel. Tras 10 volúmenes de lavado, la proteína de interés se eluyó con un gradiente lineal de imidazol, al aplicar de forma continuada el tampón de elución B (línea verde), expuesto en la Tabla III-C. La absorbancia a una longitud de onda de 280 nm (línea azul) de las proteínas

eluidas, se muestra en el cromatograma. Las muestras obtenidas se analizaron mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 17 %, TRIS-HCl. Los carriles expuestos en el mismo hacen referencia a: “Li”: Lisado celular; “I” y “S”: fracciones insoluble y soluble respectivamente tras la lisis celular y centrifugación; “NR”: muestra no retenida en la columna; “L”: lavado; “M”: marcador de alto peso molecular (*BioRad*[®]), en kDa (de los dos pesos limitantes de la banda de interés); “A4-A12”: fracciones eluidas (delimitadas en rojo en el cromatograma). El tamaño de la proteína de fusión His-SMT3-BRMS1₅₁₋₈₄ es de aproximadamente 18,2 kDa (mostrada con una flecha).

3.2.2 Proteolisis y eliminación de la cola de histidinas por columna de afinidad

Las muestras correspondientes a las fracciones A6-A11 fueron unidas, originando un volumen de 28 mL a una concentración de 12 mg/mL. Con el fin de eliminar el fragmento His-SMT3, se añadió a la totalidad de la muestra la proteasa ULP en una proporción en peso de ULP:BRMS1₅₁₋₈₄ de 1:1000. Este proceso se llevó a cabo dializando en un tampón 50 mM TRIS pH 7,5 y 0,15 M NaCl, durante toda la noche. El análisis mediante geles de poliacrilamida, teñido posteriormente con azul de Coomassie mostró dos moléculas bien diferenciadas: la proteína SMT3 unida a histidinas, de un tamaño aproximado de 14 kDa, y la proteína cortada (BRMS1₅₁₋₈₄) de aproximadamente 4,4 kDa (Figura 14; carril nombrado como “C”). Como se observa en la figura, prácticamente se consiguió la digestión de la totalidad de la proteína de fusión a histidinas.

Con el fin de eliminar tanto la proteína de fusión no digerida (His-SMT3-BRMS1₅₁₋₈₄), como el fragmento correspondiente a His-SMT3 y la enzima usada para la proteolisis que presenta una fusión a histidinas, se pasó la mezcla proteica por una columna de afinidad por histidinas cargada previamente con níquel. Tras el paso por la columna, el polipéptido BRMS1₅₁₋₈₄ libre de histidinas no se retuvo en la columna (“NR” en la Figura 14). El fragmento liberado His-SMT3 fue posteriormente eluido con imidazol.

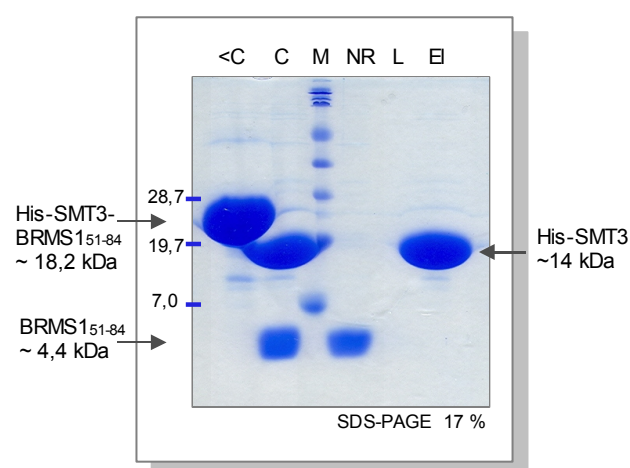


Figura 14. Muestras resultantes de la proteolisis de His-SMT3-BRMS1₅₁₋₈₄ analizadas mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 17% TRIS-HCl.

Los carriles “<C” y “C” se corresponden con muestras tomadas de la mezcla proteica antes y después de la proteolisis respectivamente. “NR” hace mención a la muestra no retenida en la columna de afinidad (BRMS1₅₁₋₈₄), y “L” al lavado realizado antes de la elución (“E”). Los marcadores de tamaño molecular (M), se muestran, en unidades de kDa, a la izquierda del gel.

His-SMT3-BRMS1₅₁₋₈₄ es la proteína inicial sin tratar con proteasa y His-SMT3 es el fragmento que estaba fusionado a la construcción BRMS1₅₁₋₈₄, los cuales han sido separados tras la proteolisis.

El polipéptido BRMS1₅₁₋₈₄ libre de cola de His-SMT3, a una concentración de 3 mg/mL en un volumen de aproximadamente 30 mL, fue concentrado hasta un volumen de

1,7 mL o 4,5 mL para ser cargado en un bucle de inyección de 2 mL o 5 mL de capacidad respectivamente, de una columna “Superdex 75 16/60 ó 26/60” de filtración en gel.

3.2.3 Cromatografía de exclusión por tamaño molecular

Tras inyectar la muestra (“<S75” en la Figura 15) en la columna de exclusión por tamaño molecular pre-equilibrada con un tampón que contenía 20 mM TRIS pH 7,5 y 50 mM NaCl, las proteínas eluyeron según su tamaño en diferentes fracciones, como se observa en la Figura 15, donde se muestra la señal de absorbancia a una longitud de onda de 280 nm. Las muestras obtenidas se analizaron en un gel SDS-PAGE teñido con solución de azul de Coomassie (Figura 15-A).

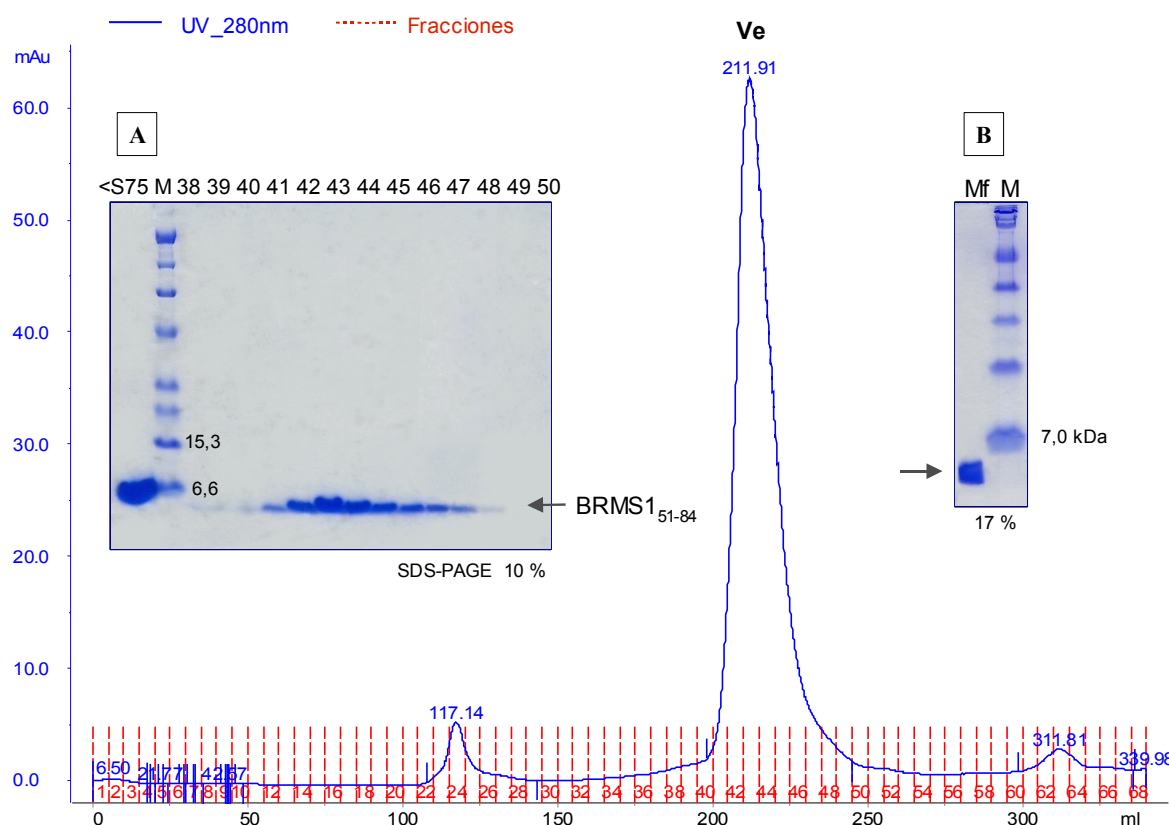


Figura 15. Cromatografía de exclusión por tamaño molecular del polipéptido BRMS1₅₁₋₈₄. El cromatograma muestra la absorbancia a una longitud de onda de 280 nm (azul), de las proteínas eluidas a distintos volúmenes según su tamaño y recogidas en fracciones de 5 mL representadas en rojo. Se muestra sobre cada pico de elución, su correspondiente volumen, siendo “Ve” el de la proteína de interés. Las muestras antes de la inyección en la columna (“<S75”) y eluidas en el pico mayoritario (38-50) se muestran en el gel de poliacrilamida (SDS-PAGE) BIS-TRIS del 10 % (*Invitrogen*^L), nombrado como gel A. Las fracciones 41 a la 46 fueron concentradas, resultando una proteína destinada a ensayos de cristalización, con un elevado grado de pureza (“Mf”), expuesta en el gel B, SDS-PAGE al 17 %, TRIS-HCl. En ambos geles se expone el marcador usado, de alto peso molecular “M” (*BioRad*^E) y se representa el tamaño en kDa de los pesos cercanos a la banda de interés. Las bandas señaladas con una flecha, cuyo peso molecular es de aproximadamente 4,4 kDa, corresponden a la proteína BRMS1₅₁₋₈₄.

Con el fin de conocer, de forma aproximada, el tamaño molecular de la proteína eluída y por tanto su conformación oligomérica más probable, los volúmenes de elución obtenidos tras la cromatografía se interpolaron en una curva de calibración obtenida previamente con estándares de tamaño molecular conocidos. El tamaño correspondiente a 211,91 mL, volumen de elución de BRMS1₅₁₋₈₄, fue de aproximadamente 9,2 kDa. Estos datos sugieren que la conformación oligomérica del polipéptido podría corresponder con la dimérica.

Las fracciones 41 a 46, correspondientes al máximo de absorción a 280 nm, fueron agrupadas obteniéndose un volumen final de 30 mL y una concentración de 2,5 mg/mL. Tras concentrar la muestra hasta un volumen de 400 µL, se obtuvo una concentración final de 150 mg/mL, en el tampón de carrera de la columna de filtración en gel (20 mM TRIS pH 7,5; 50 mM NaCl). Este tampón se complementó con 2 mM DTT cuando se destinaba la proteína final a la formación de complejo con construcciones de NMI, con el fin de mantener la muestra reducida y evitar la formación de puentes de disulfuro entre los residuos de cisteína de las moléculas de BRMS1.

La muestra final, presentaba un elevado grado de pureza, como se puede observar en el gel B, teñido con solución de Coomassie, de la Figura 15. Su identidad fue verificada como BRMS1₅₁₋₈₄ mediante análisis de espectrometría de masas.

Con el fin de caracterizar biofísicamente la proteína BRMS1₅₁₋₈₄ en solución, se realizaron ensayos de Dicroísmo Circular, utilizando un espectropolarímetro. Para ello se obtuvieron mediciones de las diferencias en absorción sobre la dirección en la que las moléculas polarizan la luz (derecha o izquierda), lo que fue usado para estimar el promedio de estructura secundaria de BRMS1₅₁₋₈₄. El experimento se realizó intentando simular, en la medida de lo posible, las condiciones de los futuros ensayos de cristalización; esto es, a 20 °C de temperatura y en un tampón 10 mM HEPES pH 7,5 y 50 mM NaCl. El espectro obtenido mostró dos mínimos de absorción en las longitudes de onda de aproximadamente 208 nm y de 222 nm (Figura 16), patrón que es característico de conformación helicoidal. Esta disposición mayoritaria promediada en α -hélice, confirma la predicción de estructura helicoidal como la de los motivos “coiled-coil”.

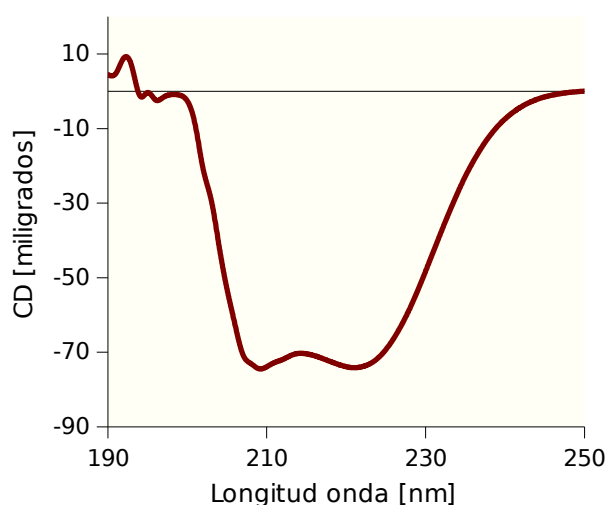


Figura 16. Espectroscopía de Dicroísmo Circular del polipéptido BRMS1₅₁₋₈₄.

Se muestra la elipticidad (o unidad de CD en miligrados) obtenida usando una concentración de proteína de 0,25 mg/mL en un tampón 10 mM HEPES pH 7,5 y 50 mM NaCl, y a 20 °C. El rango de longitud de onda medido es de 190 nm a 250 nm. Se distinguen dos mínimos a 208 y 222 nm, que coinciden con el patrón característico de α -hélice.

Una vez caracterizado biofísicamente, el polipéptido BRMS1₅₁₋₈₄ en solución fue destinado a ensayos de cristalización.

3.3 Cristalización de BRMS1₅₁₋₈₄

3.3.1 Cristalización inicial

El polipéptido BRMS1₅₁₋₈₄, en un tampón 20 mM TRIS pH 7,5 y 50 mM NaCl, fue usado para probar mas de 1000 condiciones comerciales distintas de cristalización, mediante la técnica de difusión de vapor, usando gota sentada frente a un volumen de reservorio de 50 μ L. Para su optimización, se variaron diferentes parámetros, como son la proporción entre proteína:reservorio, la concentración de proteína, la temperatura, etc.

Finalmente se obtuvieron cristales de proteína a 22 °C, en las condiciones 0,1 M Acetato sódico pH 4,6; 0,2 M Sulfato amónico y 12 % PEG 4000, tras 1 semana, en una segunda condición 0,1 M HEPES pH 7,5; 0,2 M Cloruro de magnesio y 30 % PEG 400, tras 1 día, y en 0,1 M HEPES pH 7,5; 20 % Isopropanol y 10 % PEG 4000 transcurridos 12 días. Las concentraciones del polipéptido utilizadas fueron de 85 mg/mL, 53 mg/mL y 120 mg/mL respectivamente (cristales expuestos en la Figura 17-A y -B).


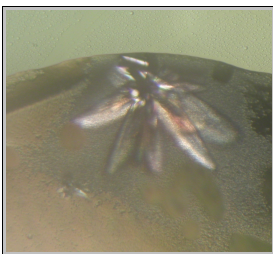
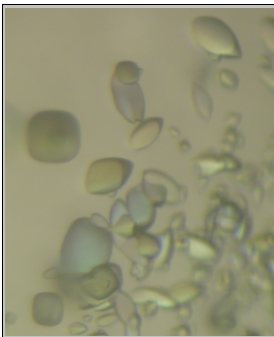
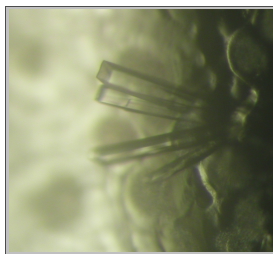
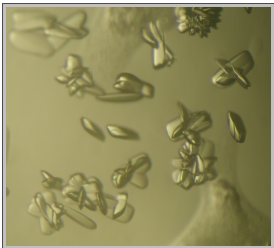
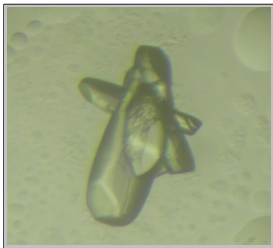
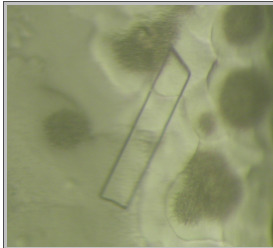
A		B	C
			
			
- 0,1 M Acetato sódico pH 4,6 - 0,2 M Sulfato amónico - 12 % PEG 4000		- 0,1 M HEPES pH 7,5 - 0,2 M Cloruro de magnesio - 30 % PEG 400	- 0,1 M HEPES pH 7.5 - 20 % Isopropanol - 10 % PEG 4000

Figura 17. Imágenes de cristales iniciales de BRMS1₅₁₋₈₄ obtenidos mediante la técnica de difusión de vapor en gota sentada, en formato de 96 pocillos. Los cristales se obtuvieron tras mezclar volúmenes iguales de proteína y reservorio, a una temperatura de 22 °C. **A**, **B** y **C** hacen referencia a las condiciones de cristalización, cuyos componentes se exponen.

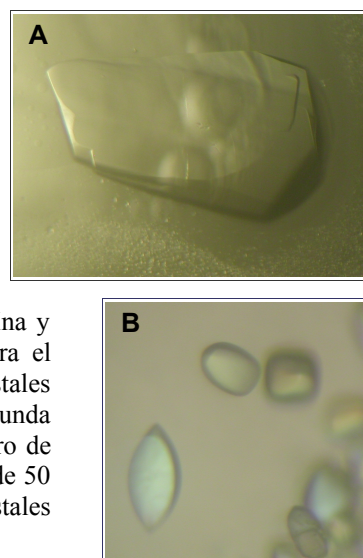
3.3.2 Optimización de cristales

Con el fin de conseguir cristales de mejor calidad que los obtenidos inicialmente, es decir, únicos y de mayor tamaño, se realizaron diversos ensayos de optimización, expuestos en la sección de *Materiales y Métodos / Optimización de cristales*, para lo que se usaron placas de 24 pocillos, utilizando el método de gota colgante.

Cada condición fue optimizada modificando condiciones como la concentración de proteína, así como el pH y la concentración de precipitante. Estos dos últimos se variaron a lo largo de cada eje de la placa, creando así un gradiente diagonal. Se consiguieron reproducir con éxito los cristales mostrados en la Figura 17-A y 17-B. Sin embargo, los cristales obtenidos en la condición de cristalización “C” no pudieron reproducirse a pesar de los numerosos intentos.

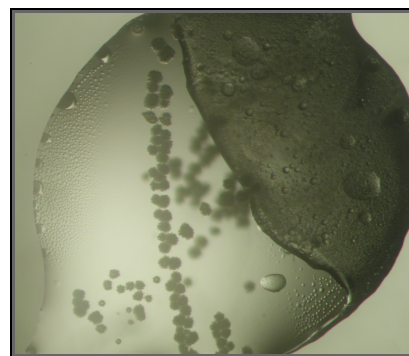
Finalmente se obtuvieron cristales optimizados modificando dos condiciones. La primera de ellas, 0,1 M Acetato sódico pH 4,3; 0,2 M Sulfato amónico y 2 % PEG 4000, con una concentración de proteína de 150 mg/mL y a diferentes proporciones de proteína:reservorio (1:1 y 1:2), permitió obtener unos cristales de unas dimensiones de 0,5 x 0,4 x 0,15 mm aproximadamente, como el que se muestra en la Figura 18-A. La segunda condición optimizada, donde se lograron los cristales expuestos en la Figura 18-B, contiene 0,1 M HEPES pH 7,0; 0,2 M Cloruro de magnesio y 20 % PEG 400. En este caso, la concentración usada fue de 50 mg/mL y bajo una proporción 1:1 proteína:reservorio. Los cristales obtenidos eran de un tamaño aproximado de 0,1 x 0,05 x 0,05 mm.

Figura 18. Cristales de BRMS1₅₁₋₈₄ optimizados obtenidos mediante la técnica de difusión de vapor en gota colgante, usando placas de 24 pocillos. Se exponen imágenes de cristales optimizados con respecto a la condición comercial expuesta en la Figura 17-A y 17-B. Los cristales nombrados como **A** corresponden a la condición 0,1 M Acetato sódico pH 4,3; 0,2 M Sulfato amónico; 2 % PEG 4000, con una concentración de proteína y proporciones entre proteína y reservorio (p:r) óptimos para el crecimiento de 150 mg/mL y 1:1 y 1:2 (la calidad de los cristales resultó la misma en ambas proporciones). En la segunda condición optimizada (0,1 M HEPES pH 7,0; 0,2 M Cloruro de magnesio; 20 % PEG 400) se utilizaron una concentración de 50 mg/mL y una proporción proteína:reservorio de 1:1. Los cristales expuestos fueron destinados a ensayos de difracción.



Es preciso mencionar que aunque intentamos optimizar aún mas los cristales (en “A” y “B”) y reproducirlos (en “C”) mediante el sembrado de centros de nucleación en una nueva gota (“seeding”), sólo surgieron cristales nuevos a lo largo de la ruta trazada con el pelo de gato bajo la condición “A”, tal y como se puede observar en un ejemplo mostrado en la Figura 19. A pesar de conseguir cristales a partir de los núcleos depositados, no se consiguieron mejorar con respecto a los ya optimizados, expuestos anteriormente.

Figura 19. Cristales de BRMS1₅₁₋₈₄ crecidos a lo largo de la ruta trazada con un pelo de gato, mediante la técnica de sembrado de centros de nucleación o “microseeding”. La condición de cristalización contiene 0,1 M Acetato sódico pH 4,6; 0,2 M Sulfato amónico y 3 % PEG 4000, obtenidos mediante la técnica de difusión de vapor en gota sentada y formato de 96 pocillos.



3.4 Resolución de la estructura de BRMS1₅₁₋₈₄

Los cristales de BRMS1₅₁₋₈₄ fueron rápidamente enfriados bajo un flujo de gas de nitrógeno y mantenidos a una temperatura de 130 K en el goniómetro del difractor, para ser irradiados con un haz de rayos X. Para ello previamente se bañaron los cristales en una gota de Paratone-N, sustituyéndose el solvente y tampón de cristalización de la superficie del cristal y del lazo por un ambiente crioprotector. La elección de esta sustancia como agente protector, se realizó tras probar otras como etilenglicol, glicerol, sacarosa o diferentes tipos de polietilenglicol a distintas concentraciones, y resultar dañinas para el cristal, tanto macroscópica (rotura del cristal) como microscópicamente (daño en el patrón de difracción).

3.4.1 Recolección de datos de difracción del cristal nativo de BRMS1₅₁₋₈₄

Debido a la diferencia de tamaño, los cristales se montaron en lazos de diferente diámetro, siendo de 0,3-0,4 mm para el cristal mostrado en la Figura 18-A, y de 0,1-0,2 mm para los de la Figura 18-B.

La difracción de ambos cristales fue probada en la fuente de rayos de X de la línea ID14-1 del sincrotrón de Grenoble, utilizando una longitud de onda de 0,934 Å. Los cristales obtenidos bajo la condición “B” no difractaron, a diferencia del cristal “A”, que lo hizo a una resolución de 2,0 Å.

Con el fin de obtener unos datos mas completos e informativos, y evitar reflexiones sobresaturadas en el área de baja resolución, se recogieron dos juegos de datos limitando diferentes rangos de resolución. Un primer juego de datos recoge las señales difractadas por el cristal a alta resolución (AR en la Tabla VIII), hasta 2,0 Å, y un segundo juego de datos a baja (BR en la Tabla VIII), limitando la resolución máxima a 5,0 Å. Estos dos juegos de datos fueron recogidos sobre el mismo cristal. Los parámetros usados en la recolección de los mismos se exponen en la Tabla VIII.

Juego de datos	Energía	Distancia	Resolución	Incremento de ϕ total	Incremento de ϕ / Imagen	Tiempo de exposición	Número de pases
AR	0,934 Å (13,274 keV)	229,63 mm	57,544 - 2,000 Å	180 °	1 °	2 seg.	6
BR	0,934 Å (13,274 keV)	551,82 mm	95,783 - 5,000 Å	360 °	2 °	0,4 seg.	3

Tabla VIII. Parámetros usados en la recolección de los juegos de datos a alta y baja resolución (**AR** y **BR** respectivamente) de un cristal de BRMS1₅₁₋₈₄ en la línea ID14-1 del sincrotrón. Se muestran por orden, la longitud de onda y energía con la que se emite el haz de rayos X incidente sobre el cristal; la distancia entre el cristal y el detector y los límites de resolución a la que se recogieron las difracciones emitidas por las moléculas del cristal. A continuación se muestran parámetros del ángulo “phi” (ϕ) que determina la orientación del cristal donde se soporta el cristal. Estos se refieren al incremento total del ángulo recolectado y al incremento en la rotación del mismo en cada imagen. También se expone el tiempo de exposición en segundos para la toma de cada imagen y el número de pases realizados sobre cada imagen.

El patrón de difracción y los límites de resolución en las imágenes difractadas para los dos juegos de datos recogidos, se pueden ver en la Figura 20.

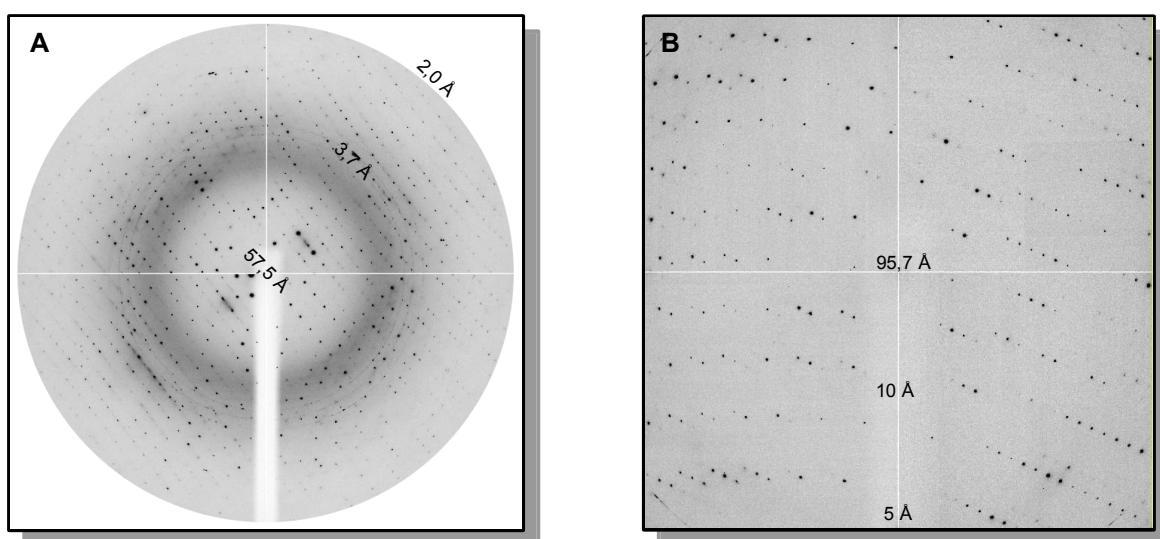


Figura 20. Imágenes representativas de espectros de difracción de los juegos de datos recogidos a alta (**A**) y a baja (**B**) resolución de un cristal de BRMS1₅₁₋₈₄, obtenidas en la línea ID14-1 del sincrotrón. Sobre estas figuras se muestran algunos límites de resolución, expresados en angstroms (Å). Las imágenes se corresponden a incrementos en el ángulo ϕ de 1 y 2 °, durante 2 y 0,4 seg. Respectivamente, y fueron recogidas a una distancia entre el cristal y el detector de 230 mm para la alta y de 552 mm para la baja resolución.

3.4.2 *Recolección de datos de difracción del cristal derivado de mercurio de BRMS1₅₁₋₈₄*

Con el fin de intentar obtener las fases de las ondas que conforman cada reflexión del juego de datos, abordamos dos posibilidades. En un primer lugar, intentamos realizar reemplazo molecular con alguna estructura registrada en la base de datos PDB con

secuencia similar a BRMS1. Esta alternativa quedó descartada tras no encontrar un modelo conservado en cuanto a la secuencia, y tras probar, sin éxito, reemplazo molecular con numerosos tipos de “coiled-coil”. Una segunda posibilidad se centró en la consecución de un cristal derivado con átomos pesados, que pudiese ser usado para generar señal anómala y/o isomorfa sobre el patrón de difracción y orientar con ello, todos los átomos del cristal. Para esta segunda alternativa se intentó en primer lugar obtener cristales de proteína con sustitución de la metionina por seleno-metionina, obtenida según el protocolo explicado en *Materiales y Métodos*, pero no se obtuvieron cristales, por lo que decidimos abordar la obtención de estructuras derivadas mediante impregnación de los cristales ya formados en soluciones conteniendo sales de átomos pesados.

La afinidad de cada compuesto con átomos pesados por sitios concretos de la proteína en el cristal es específica, dependiendo tanto de la secuencia aminoacídica de la molécula, es decir de su química, como de la forma y empaquetamiento del cristal. Esto explica que unos átomos pesados pueden incorporarse correctamente dentro de una estructura cristalina y originar buena señal anómala y otros no. Por ello, realizamos un análisis minucioso de la secuencia de BRMS1₅₁₋₈₄, en los que observando la presencia de un residuo de cisteína (para el que son afines los elementos mercurio, oro y platino) y varios residuos positivos y una metionina (por los que el platino mantiene afinidad). Así, *a priori*, los grupos de átomos pesados que engloban 37 derivados de mercurio, platino y oro en total, serían adecuados para la interacción con el polipéptido.

Con el fin de discriminar cuál de todos presentaba experimentalmente una mayor afinidad por el polipéptido, se realizó una selección basada en la estabilidad del átomo pesado en el pH de crecimiento del cristal. Con ello se descartaron 13 de los 37, cuyo rango de estabilidad de pH estaba fuera de la condición de cristalización (*Hampton Research*^V). Los 24 derivados de sales pesadas restantes se analizaron según la diferencia en la migración en un gel nativo de la proteína incubada con dicha sal en solución. Para ello, se incubó la proteína en solución durante 30 minutos con los diferentes compuestos, preparados en la condición de cristalización, y posteriormente se estudió su movilidad electroforética en el gel nativo, teñido con azul de Coomassie (Figura 21).

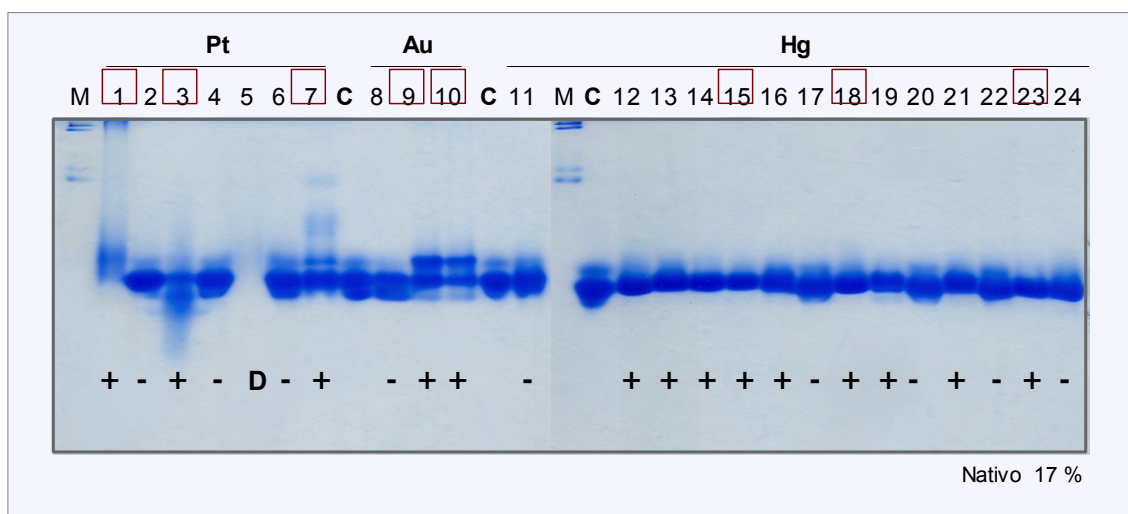


Figura 21. Electroforesis en gel de poliacrilamida al 17 % en condiciones nativas de muestras del polipéptido BRMS1₅₁₋₈₄ incubada con 24 derivados de átomos pesados de platino (Pt), oro (Au) y mercurio (Hg). Se denotan diferencias en los patrones de migración del polipéptido incubando con los diferentes

derivados de átomos pesados. **C**: carriles en los que se ha cargado la proteína control, sin incubar con metal pesado. Con un símbolo “-” se muestran los derivados que mantienen el mismo patrón de migración que la proteína control, y con un símbolo “+” aquellas que, por el contrario, sí presentan diferencias en el comportamiento en el gel. Marcados como **D** se observan las muestras que no han entrado en el carril. Con un recuadro rojo se diferencian los derivados de metales pesados llevados a la fase experimental de impregnación del cristal con los mismos. El marcador usado (**M**), se trata de una mezcla de aprotinina, ferritina A y citocromo, preparada en el laboratorio.

Pudimos diferenciar tres tipos de comportamiento según la diferencia en la migración mostrada en el gel nativo:

1.- Proteína que no cambia su patrón de movilidad con respecto al control negativo (proteína sin incubar con ningún átomo pesado; “**C**” en la Figura 21). Este comportamiento sugiere que el polipéptido no interacciona con el ligando. Por lo tanto, estos derivados (marcados con el símbolo “-” en el gel) fueron descartados.

2.- Proteína que no llega a entrar en el gel, de forma que se observa el carril vacío. Estas sales también fueron descartadas por ejercer un efecto negativo o desnaturalizante sobre la conformación o plegamiento de la proteína, y aparecen marcadas como “**D**” en la Figura 21.

3.- Por último, algunos derivados pesados de los tres elementos (Hg, Au y Pt), cambiaron claramente el patrón de migración de BRMS1₅₁₋₈₄ (señalados con un símbolo “+” en la Figura 21), por ello fueron seleccionados y utilizados en ensayos posteriores. Estas 8 sales de átomo pesado mostradas con un recuadro rojo en el gel nativo de la Figura 21, se detallan a continuación (Tabla IX).

Pt			Au		Hg		
1	3	7	9	10	15	18	23
K ₂ PtCl ₄	K ₂ Pt(NO ₂) ₄	[Pt ₂ I ₂ (H ₂ NCH ₂ CH ₂ NH ₂) ₂](NO ₃) ₂	KAuCl ₄	AuCl ₃	C ₈ H ₈ HgO ₂	C ₂ H ₅ HgCl	Hg(NO ₃) ₂ ·H ₂ O

Tabla IX. Derivados de átomos pesados de platino, oro y mercurio, seleccionados para impregnar los cristales de BRMS1₅₁₋₈₄. Estas sales mostraron afinidad por el polipéptido, según se mostró en el comportamiento analizado en el gel nativo de la Figura 21. La numeración se corresponde con la utilizada en dicho gel.

Los átomos pesados seleccionados fueron usados para su incorporación a cristales de BRMS1₅₁₋₈₄ obtenidos previamente, como los que se muestran en la Figura 22 mediante baño o impregnación de los mismos en dichos derivados (en inglés “soaking”). Se probaron diferentes combinaciones de concentraciones y tiempos de exposición. Normalmente, para concentraciones mayores (en torno a 5-10 mM), se usaron tiempos menores (5 min - 1 hora), y para concentraciones menores (1-2 mM), se incubaron durante mayor tiempo (1 h, 2 h o toda la noche). Esta decisión se basó según la estabilidad que el cristal parecía tener a nivel macroscópico.

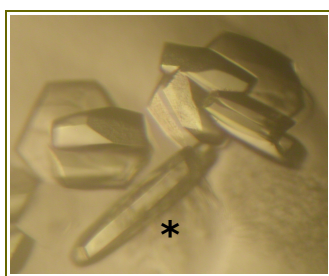


Figura 22. Ejemplo de cristales de BRMS1₅₁₋₈₄ mejorados usados para su incubación en soluciones de sales de átomos pesados de platino, oro y mercurio (“soaking”) y obtención de cristal derivado. El cristal marcado con un asterisco fue usado para obtener un juego de datos de imágenes de difracción de un cristal derivado de nitrato de mercurio monohidratado.

Tras análisis de difracción se pudo comprobar como varias de estas 8 sales, concretamente las números 1, 15 y 18 en la Tabla IX, modificaban la red cristalina, afectando macroscópicamente al cristal o a su capacidad de difracción. De las restantes, se recogieron 5 juegos de datos, correspondientes a los derivados números 3, 7, 9, 10 y 23 de la Tabla IX. Estos datos fueron indexados e integrados y posteriormente escalados, utilizando los programas *Mosflm* y *Scala* de *CCP4* repectivamente. Sin embargo, únicamente del derivado número 23 (nitrato de mercurio (II) monohidratado), se pudo obtener buena y suficiente señal anómala que nos sirvió para obtener las fases. La concentración y tiempo de exposición óptimos del cristal en la solución de nitrato de mercurio fue de 2 mM, durante 15 horas.

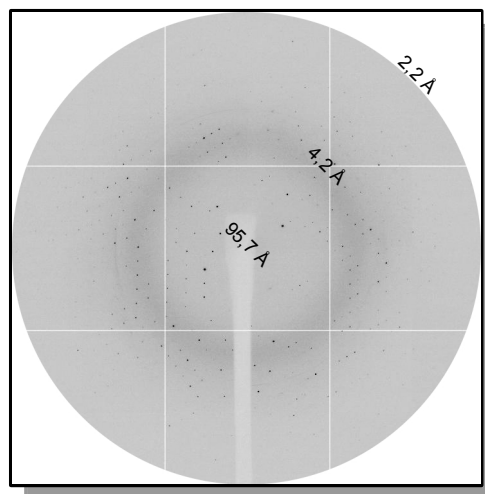
El cristal derivado de nitrato de mercurio (marcado con un asterisco en la Figura 22) se montó en un lazo de 0,05-0,1 mm de tamaño y fue enfriado con gas de nitrógeno a una temperatura de 130 K, para ser irradiado en la línea ID23-1 del sincrotrón de Grenoble, usando una longitud de onda de 0,9729 Å. Los parámetros usados en la recolección de los datos de difracción se exponen en la Tabla X.

Energía	Distancia	Resolución	Incremento de ϕ total	Incremento de ϕ / Imagen	Tiempo de exposición	Nº de pases	Atenuación del rayo
0,9729 Å (12,743 keV)	322,9 mm	95,783 - 2,200 Å	180 °	1 °	0,2 seg.	1	15 %

Tabla X. Parámetros usados en la recolección del conjunto de datos de difracción de un cristal de BRMS1₅₁₋₈₄ derivado de mercurio en la línea ID23-1 del sincrotrón. Se muestran por orden, la longitud de onda y energía con la que se emite el haz de rayos X incidente sobre el cristal; la distancia entre el cristal y el detector y los límites de resolución a la que se recogieron las difracciones emitidas por las moléculas del cristal. A continuación se muestran parámetros del ángulo “phi” (ϕ) que determina el posicionamiento del goniómetro donde se soporta el cristal. Estos son el ángulos total colectado y el incremento en la rotación del mismo en cada imagen. También se expone el tiempo de exposición en segundos para la toma de cada imagen, el número de repeticiones tomadas de cada imagen y el porcentaje de atenuación del rayo.

Finalmente se recogió un total de 180 °, alcanzando una resolución de 2,2 Å. Se obtuvieron unas imágenes con un patrón de difracción como el que se expone en la Figura 23.

Figura 23. Patrón de difracción de un cristal de BRMS1₅₁₋₈₄ derivado con nitrato de mercurio, obtenido en la línea ID23-1 del sincrotrón. Se recogieron 180 °, entre un rango de resolución de 95,8 y 2,2 Å, y una distancia entre el cristal y el detector de unos 323 mm. Cada imagen se expuso 0,2 seg, rotando 1 °. Sobre la figura se muestran algunos límites de resolución, expresados en angstroms (Å).



3.4.3 Procesado de datos y estructura de BRMS1₅₁₋₈₄

Los datos de difracción obtenidos, tanto del cristal nativo como del cristal derivado de mercurio, se utilizaron para su procesamiento con diferentes programas informáticos de análisis. La reducción de datos se realizó usando el programa *Mosflm* para el indexado e integrado, y la herramienta *Scala* para el escalado, ambos programas del paquete *CCP4*. Tras indexar las reflexiones se obtuvieron las dimensiones de la celda unidad del cristal; esto es, la longitud de sus ejes y los ángulos que la forman (expuestas en la Tabla XI), al igual que la intensidad integrada de cada reflexión única, el error experimental en la medida y el grupo espacial de cada cristal, tratándose en ambos casos de $C222_1$, de la familia rectangular-ortorrómbico y con la celda centrada en la cara *C*.

Los dos juegos de datos recogidos del cristal nativo a alta y baja resolución (AR y BR), fueron indexados e integrados de forma individual y posteriormente combinados y escalados juntos. Por ello, nos referiremos a partir de ahora a un único juego de datos nativo, cuyos parámetros se exponen en la Tabla XI, junto a los del procesamiento del derivado.

BRMS1 aa (51-84)		
Reducción de datos	Nativo	Derivado (Hg)
Grupo espacial	C222 ₁	C222 ₁
Dimensiones de la celda		
a; b; c (Å)	42,636; 191,274; 71,940	42,414; 191,403; 71,868
α , β , γ (°)	90; 90; 90	90; 90; 90
Rango de resolución (Å) (última capa)	95,78 - 2,00 (2,11 - 2,00)	95,78 - 2,20 (2,32 - 2,20)
Observaciones (únicas)	146014 (20344)	108281 (15378)
Completitud (%) (última capa)	99,2 (94,7)	100,0 (100,0)
Multiplicidad (última capa)	7,2 (3,9)	7,0 (7,2)
R _{sym} ^a (última capa)	0,052 (0,202)	0,088 (0,496)
R _{meas} ^b (última capa)	0,055 (0,237)	0,106 (0,577)
R _{pim} ^c (última capa)	0,019 (0,119)	0,041 (0,214)
$\langle I/\sigma \rangle$ (última capa)	28,2 (7,2)	15,4 (3,7)
Completitud anómala (%)		99,9
Multiplicidad anómala		3,7
DelAnom correlation between half-sets		0,212
Mid-Slope of Anom Normal Probability		1,168

^a R_{sym} es el *factor R* no sopesado sobre las intensidades simétricamente relacionadas.

$R_{sym} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, donde $I_i(hkl)$ es la intensidad de la reflexión hkl y $\langle I(hkl) \rangle$ es la intensidad media de todas las observaciones i de la reflexión hkl .

^b $R_{meas} = \{\sum_{hkl} [N/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|\} / \sum_{hkl} \sum_i I_i(hkl)$, donde $I_i(hkl)$ son las intensidades observadas, $\langle I(hkl) \rangle$ son las intensidades medias y N es la multiplicidad de la reflexión hkl .

^c $R_{pim} = \{\sum_{hkl} [1/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|\} / \sum_{hkl} \sum_i I_i(hkl)$, donde $I_i(hkl)$ son las intensidades observadas, $\langle I(hkl) \rangle$ son las intensidades medias y N es la multiplicidad de la reflexión hkl .

$\langle I/\sigma \rangle$ es la media de la relación entre la intensidad de la difracción y el ruido de fondo.

Tabla XI. Datos estadísticos cristalográficos obtenidos tras el procesamiento y simplificación de los datos de difracción recogidos de los cristales nativo y derivado de mercurio de la proteína BRMS1₅₁₋₈₄.

La estimación del número de moléculas que componen la unidad asimétrica, así como el porcentaje de solvente, se realizó calculando el coeficiente de *Matthews* dentro del paquete *CCP4*. Este coeficiente calcula las probabilidades normalizadas para cada posible número de subunidades en la unidad asimétrica, utilizando para ello las dimensiones de la celda unidad, el grupo espacial y el tamaño en unidades de Dalton (*Da*) o número de residuos de la proteína. La resolución puede también ser usada para afinar la estimación de la distribución de probabilidad (Kantardjieff y Rupp, 2003; Matthews, 1968). Con ello, de entre todas las opciones, las probabilidades mas elevadas dadas por el programa, y acotando el coeficiente de *Matthews* entre 2 y 3 Å³/Da, sugieren la presencia de entre 6 y 8 moléculas en la unidad asimétrica y un contenido en solvente de entre 43 % y 57 %. Estos datos se muestran en la Tabla XII.

Nº moléculas / UA	Coefficiente de Matthews (Å ³ /Da)	Solvente (%)	P(2,00)	P(Total)
5	3,36	63,40	0,04	0,06
6	2,80	56,08	0,17	0,21
7	2,40	48,76	0,38	0,38
8	2,10	41,44	0,34	0,28
9	1,87	34,12	0,07	0,05

Tabla XII. Posibles contenidos de moléculas y solvente en la unidad asimétrica (UA) para el cristal nativo de BRMS1₅₁₋₈₄ estimados mediante cálculo de Coeficientes de *Matthews* (*CCP4*) (Kantardjieff y Rupp, 2003; Matthews, 1968), que sugiere entre 6 y 8 moléculas. P(2,00) alude a la probabilidad usando el límite de resolución de 2 Å, y P(Total) a la probabilidad a través de todos los rangos de resolución.

El mercurio incorporado al cristal aportó perturbaciones en la difracción, generando diferencias isomorfas (sobre la intensidad recogida de cada reflexión) y señal anómala (donde se da un cambio de intensidad medible, generada por la absorción de la radiación del átomo pesado) suficientes para obtener las fases. La medición de estas diferencias se hace en base a las diferencias en amplitudes de los factores entre los dos juegos de datos, que se combinaron para calcular un mapa de Patterson (en *CCP4*), en el cual se observaron picos claros mostrando 5 posiciones de átomos pesados. Un ejemplo de dos de estos picos se muestran en la Figura 24.

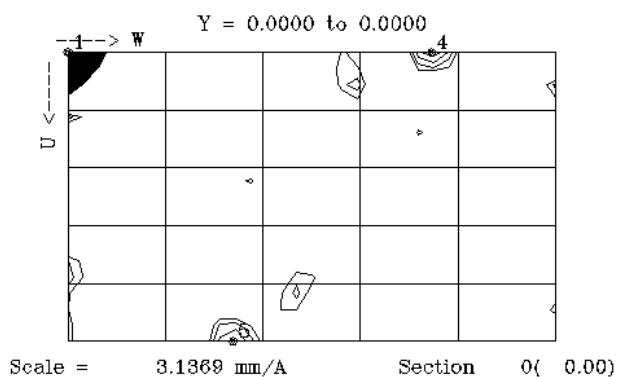


Figura 24. Ejemplo de sección de Harker de un mapa de *Patterson* de diferencias para un cristal de BRMS1₅₁₋₈₄ derivado de mercurio. En la figura se observan dos picos (numerados como 4 y 5), correspondientes a las posiciones de los átomos de mercurio en la unidad asimétrica.

Los cristales nativo y derivado resultaron ser isomorfos, con diferencias de menos de 0,25 Å entre las longitudes de los ejes de la celda. Por lo tanto, se pudieron medir las diferencias isomorfas y anómalas, lo que nos permitió aplicar la técnica de SIRAS para resolver las fases. El mejor contraste obtenido para la obtención de las fases, fue asumiendo un contenido de solvente de 0,6 (60 %); usando los datos comprendidos hasta una resolución de 3,5 Å (en *SHELXD*) y realizando 20 ciclos de faseado (en *SHELXE*).

Tras el procesamiento con *SHELXD*, de igual manera que los análisis previamente expuestos realizados con *CCP4* para calcular un mapa de *Patterson*, se encontraron cinco posiciones de mercurio por unidad asimétrica, con ocupaciones muy diversas, de modo

que para el átomo de Hg 1 era de 100 %, 89 % para el átomo de Hg 2 y 35 %, 14 % y 1 % para las posiciones de mercurio 3, 4 y 5 respectivamente.

Las posiciones de los átomos pesados, nos permitió obtener un mapa de densidad electrónica experimental inicial, observado en la Figura 25, que nos confirma que las fases se asignaron correctamente. En este mapa se intuyen claramente morfologías helicoidales, lo que nos adelantó el tipo de estructura de la que se trataba.

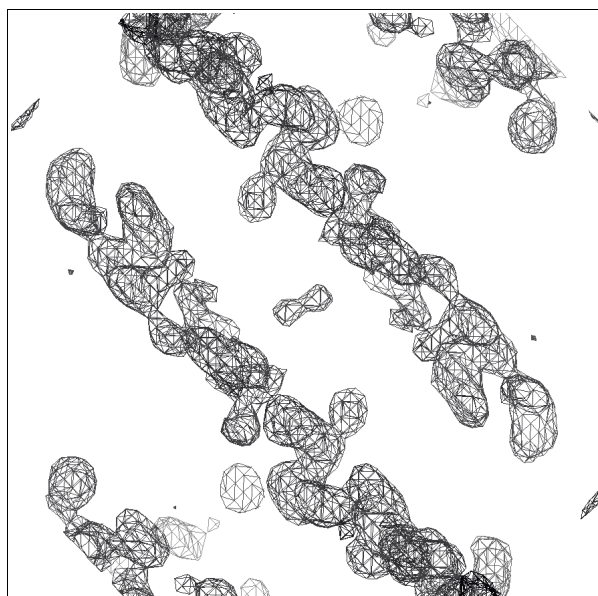


Figura 25. Mapa de densidad electrónica experimental inicial, calculado por *Coot* tras asignar las fases usando la técnica SIRAS, para la resolución de la estructura BRMS1₅₁₋₈₄. Mapa representado a un nivel de $1,5 \sigma$.

Observando el mapa experimental trazado alrededor de estos átomos (Figura 26), se puede comprobar que el átomo de mercurio número 4 se localiza en una zona que probablemente esté ocupada por el modelo BRMS1₅₁₋₈₄, y el átomo número 5 claramente no está justificado por ninguna densidad electrónica.

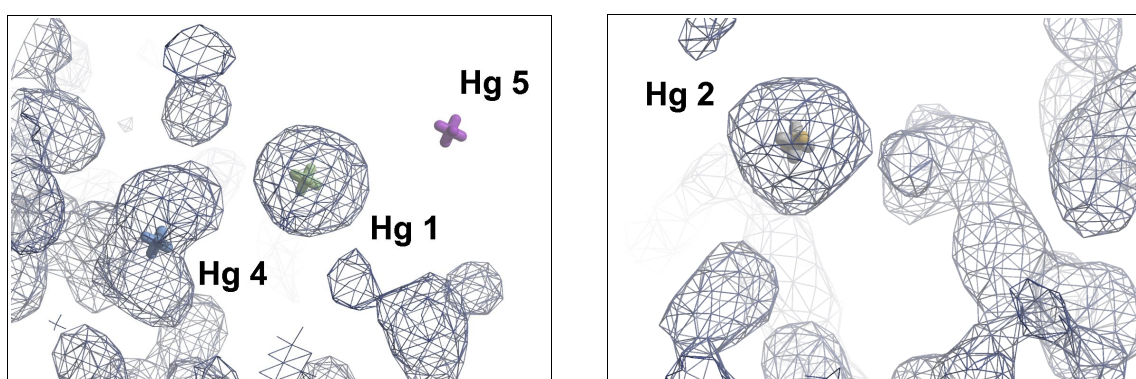


Figura 26. Mapa experimental de densidad electrónica a $1,5 \sigma$; en torno a cuatro átomos de mercurio (cruces) localizados mediante el programa *SHELXD*. Los átomos de Hg 1 y 2 tienen una ocupación de mas de un 80 %, mientras que los tres restantes de menos de 35 %. Se muestran también como ejemplos los átomos de mercurio 4, que se posiciona en la densidad electrónica correspondiente al modelo, y el 5, que no está explicado por mapa experimental.

Una vez obtenidas las fases, y usando únicamente las posiciones de los átomos de mercurio con una ocupación mayor del 80 % (Hg 1 y Hg 2), se probaron diversas técnicas para intentar construir un modelo sobre el mapa de densidad electrónica obtenido. Finalmente, combinando los modelos iniciales construidos con los programas *ARP/wARP* y *Buccaneer* (del paquete *CCP4*) y *Resolve* de *PHENIX* (*AutoBuild wizard*) se obtuvo un modelo inicial aunque mas completo. Se utilizó el juego de datos nativo (que difractó a mayor resolución), para extender la resolución de los datos utilizados en el afinamiento y así optimizar el mapa de densidad electrónica. Para ello, se afinó el modelo inicial usando las reflexiones escaladas del juego de datos nativo (con *Refmac5*). En el esquema representado en la Figura 27 se muestra, de forma resumida, la estrategia llevada a cabo para resolver la estructura, basándose en la resolución de las fases utilizando la señal anómala y las diferencias isomorfas originadas por un derivado de átomo pesado (*SIRAS*).

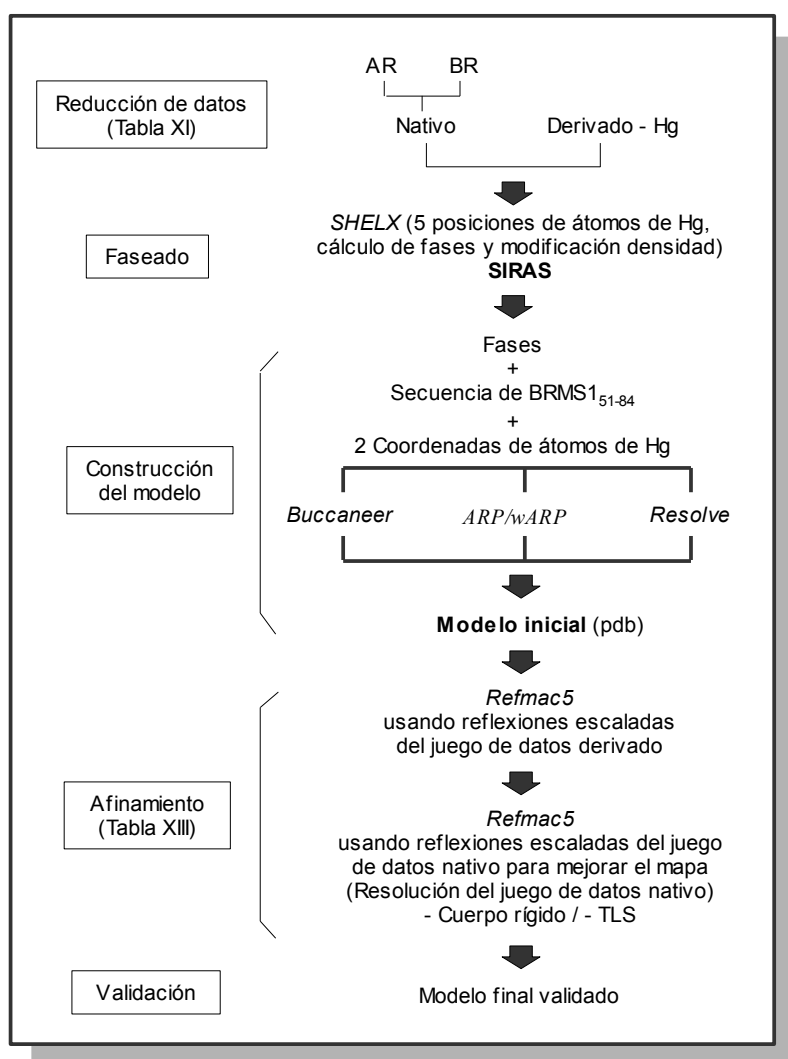


Figura 27. Estrategia utilizada para resolver la estructura de *BRMS1*₅₁₋₈₄ mediante *SIRAS*.

En cursiva se muestran los programas usados en cada caso.

AR y BR: juegos de datos del cristal nativo recogidos a alta y baja resolución respectivamente.

Para el proceso de afinado, se combinaron ciclos de afinamiento manual en el espacio real (*Coot*) y automático en el espacio recíproco con *Refmac5* (usando 8 grupos de *TLS* (uno por cadena) y realizando al principio un *afinamiento de cuerpo rígido*), se obtuvo un modelo final con unos parámetros, contenidos en la Tabla XIII, que dan idea de la fiabilidad del modelo trazado, llamados *factores R* (con la totalidad de las reflexiones)

y R_{free} (calculado con el 5 % de la reflexiones que no participan en el afinamiento) de 20,5 % y 26,8 % respectivamente, que marcan las diferencias relativas entre los factores de estructura observados (F_{obs}) y calculados (F_{calc}).

También se calcularon las desviaciones cuadráticas medias de las longitudes de los enlaces y de los ángulos, siendo de 0,013 Å y 1,398 Å para ambos respectivamente (con *Refmac5* de *CCP4*). Asimismo, se calcularon los *factores B* que estiman la vibración del átomo con respecto a su posición central, utilizando para ello la herramienta *Baverage* del programa de *CCP4*. Estos parámetros fueron de 31,43 Å² para la proteína, de 42,78 Å² para las 140 moléculas de agua modeladas y de 36,53 Å² para los heteroátomos, concretamente un sulfato, un cloro y un acético.

BRMS1 aa (51-84)	
Afinamiento	
Nº átomos no hidrógenos (moléculas de agua)	1987 (140)
Heteroátomos	
SO ₄	1
Cl	1
Acy	1
R^b (R_{free}^c)	0,205 (0,268)
$R.m.s.d.^d$ enlaces (Å)	0,013
$R.m.s.d.^d$ ángulos (°)	1,398
$Factor B^e$ (Å ²)	
Proteína	31,43
Solvente	42,78
Heteroátomos	36,53

^b $R = \sum_{hkl} \{ [F_{obs}(hkl)] - [F_{calc}(hkl)] \} / \sum_{hkl} [F_{obs}(hkl)]$, donde F_{obs} y F_{calc} son los factores de estructura observados y calculados respectivamente.

^c R_{free} es un parámetro de validación, correspondiente al *factor R* del 5 % de reflexiones que no han sido introducidas en el proceso de afinamiento.

^d $R.m.s.d.$ del inglés “Root mean square deviation”, es la desviación cuadrática media de las longitudes.

^e $Factor B$ es un parámetro de estimación de la vibración del átomo con respecto a su posición central.

Tabla XIII. Datos estadísticos cristalográficos obtenidos tras el proceso de afinamiento del modelo obtenido de BRMS1₅₁₋₈₄, con *Refmac5* (*CCP4*). SO₄, Cl y Acy hacen referencia a sulfato, cloro y acético respectivamente.

Como se muestra en la Figura 28, la unidad asimétrica contiene, además de las moléculas de agua y de los heteroátomos, 8 moléculas de BRMS1₅₁₋₈₄ (*A-H*), unidas por pares (*A-D*; *B-E*; *C-H*; *F-G*), mediante un puente disulfuro a través del residuo de cisteína número 10 de cada cadena. De los 280 residuos totales (35 residuos por cadena x 8 cadenas), se pudieron modelar un total de 208, que se muestran por cadenas en la Tabla XIV. Las cadenas *A* y *C* están completas, mientras que la secuencia de las seis restantes no

se pudo completar. De los 208 residuos modelados, 12 de ellos se asignaron ocupación doble a sus cadenas laterales, con el fin de explicar mejor la densidad electrónica.

A	B	C	D	E	F	G	H
0-34	8-34	0-34	8-25	7-27	4-26	8-34	8-29

Tabla XIV. Residuos modelados de las 8 cadenas (A-H) de la unidad asimétrica del cristal de BRMS1₅₁₋₈₄. El primer número corresponde al residuo amino-terminal, y el segundo al carboxilo-terminal. En azul se muestran las dos cadenas que pudieron ser modeladas en su totalidad.

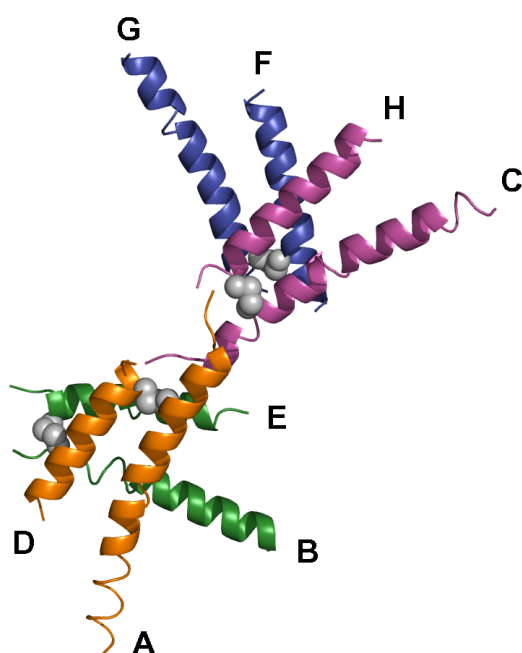
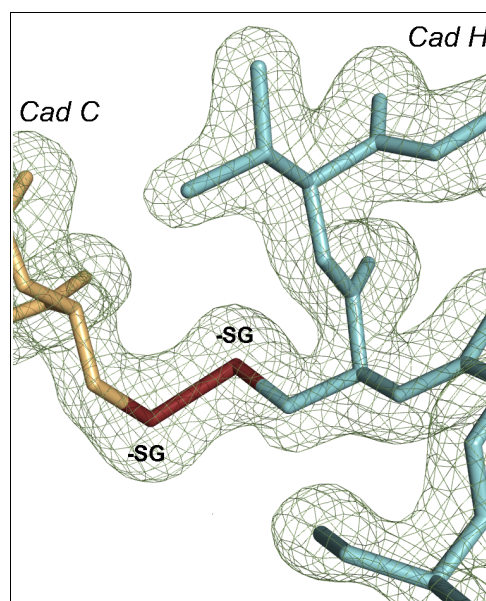


Figura 28. Unidad asimétrica conteniendo 8 moléculas (A-H) de BRMS1₅₁₋₈₄. Los pares de cadenas unidas por el puente disulfuro (esferas grises) se muestran en el mismo color. Programa *Pymol*.

El análisis detallado de los puentes disulfuro establecidos entre los cuatro pares de cadenas, concluye que la distancia media es de 2 Å entre los átomos de azufre (representados como -SG en la Figura 29) de las cisteínas que lo forman, concretamente los residuos número 10 (Cys60 en el contexto de la proteína completa) de cada cadena. El mapa de densidad electrónica 2Fo - Fc en torno a los cuatro puentes disulfuro establecidos en la unidad asimétrica, permitió inequívocamente, trazar todos los enlaces, un ejemplo de ello es el mostrado en la Figura 29.

Figura 29. Puente disulfuro establecido entre los residuos de cisteína 10 de las cadenas C (*Cad C*) y H (*Cad H*) contenidas en la unidad asimétrica del cristal de BRMS1₅₁₋₈₄. El enlace, de unos 2 Å de longitud, tiene lugar a través de los átomos -SG de ambos residuos. El mapa de densidad electrónica 2Fo - Fc está representado a un nivel de 1,5 σ . Figura realizada con *Pymol*.



Los análisis de predicción realizados sobre la proteína completa de BRMS1, mostraron, como se ha mencionado al principio de esta sección, la existencia de un motivo “coiled-coil” comprendido entre los residuos 51 a 84, que se corresponde con el polipéptido cuya estructura se ha resuelto. Con el fin de comprobar la predicción de BRMS1₅₁₋₈₄ de ser motivo “coiled-coil”, buscamos en la celdilla pares de moléculas que pudiesen responder a la geometría de estos motivos. Finalmente, localizamos dos pares de cadenas: *A-A_{sim2}* y *G-C_{sim5}* (donde *A_{sim2}* es la cadena *A* del grupo simétrico 2 y *C_{sim5}*, la cadena *C* del grupo simétrico 5), contactadas de forma levógira y antiparalela, que podían responder a este motivo.

Posteriormente analizamos los contactos establecidos entre ellas, mediante el programa Monster, que se exponen a continuación.

3.4.3.1 *Monster, análisis de interacciones entre las dos cadenas de los motivos “coiled-coil”*

Las cadenas que conforman los pares *A-A_{sim2}* y *G-C_{sim5}*, abarcan distancias comprendidas entre 2,7 Å y 6 Å. Con el fin de analizar los contactos establecidos entre las dos cadenas, y debido a que todos los residuos de la cadena *A* y *C* se pudieron modelar, pero no los de la *G*, se estudiaron únicamente los establecidos dentro del par *A-A_{sim2}*, y son estos los que se muestran en la Figura 30.

Principalmente, dichos contactos son de carácter hidrofóbico, existiendo algún puente de hidrógeno e interacciones electrostáticas entre residuos de ambas cadenas. La mayoría de los aminoácidos localizados en las posiciones *a* y *d* de la “heptada”, señalados en la Figura 30, y que mantienen contactos con otros residuos, son de carácter hidrofóbicos, como Val, Leu o Phe.

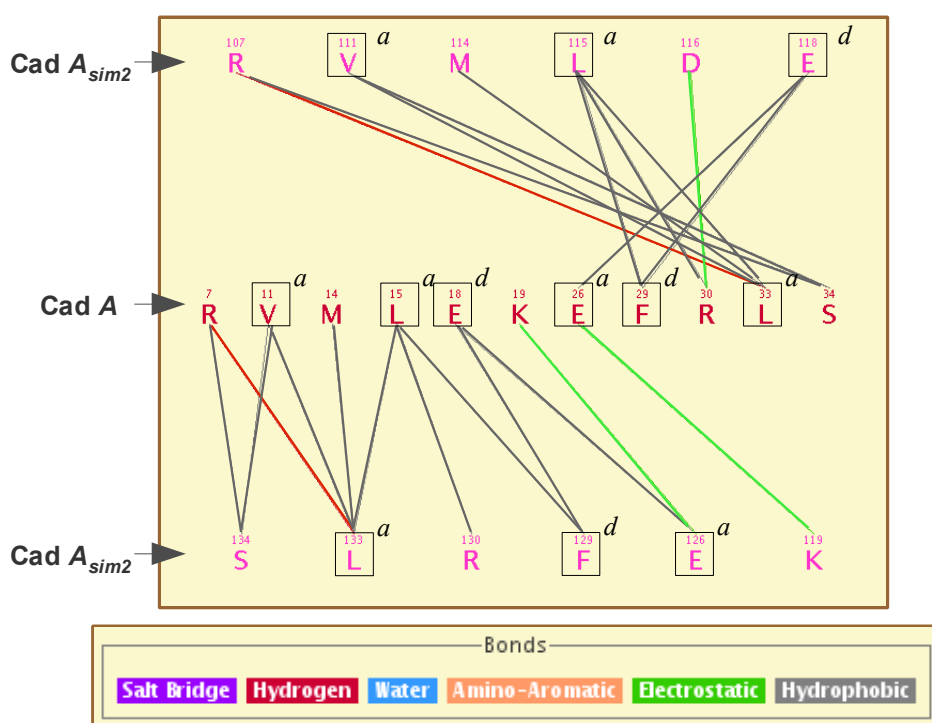


Figura 30. Contactos mantenidos entre los residuos de las dos cadenas A - A_{sim2} de BRMS1₅₁₋₈₄, analizados con el programa *Monster*. Los residuos que se localizan en las posiciones a y d de la “heptada” se muestran recuadrados. Los contactos son principalmente hidrofóbicos (color gris), existiendo algún puente de hidrógeno (rojo) y fuerzas electrostáticas (verde). Los residuos se enumeran en la parte superior de cada uno, de acuerdo a la secuencia S0-S34. La cadena A se representa en color rojo, mientras que la A_{sim2} se muestra en color rosa, con el número de residuo precedido por un “1”.

Se calcularon, mediante el programa *Contact* de *CCP4*, los parámetros de contacto entre las dos superficies, así como sus distancias medias, siendo de 1966,4 Å² la superficie total de contacto entre las cadenas A - A_{sim2} . Las distancias entre la cadenas principales (carbonos α) de las dos moléculas, varían desde 6,61 Å en la región más cercana, hasta 8,72 Å en la más lejana. En la Figura 31 se muestra la estructura del par de cadenas A - A_{sim2} . En dicha figura se representan las cadenas laterales de los residuos que forman parte de los contactos que sustentan la estructura, que como mencionamos anteriormente, son de carácter hidrofóbico, en los extremos de la doble hélice, y polares en la zona central.

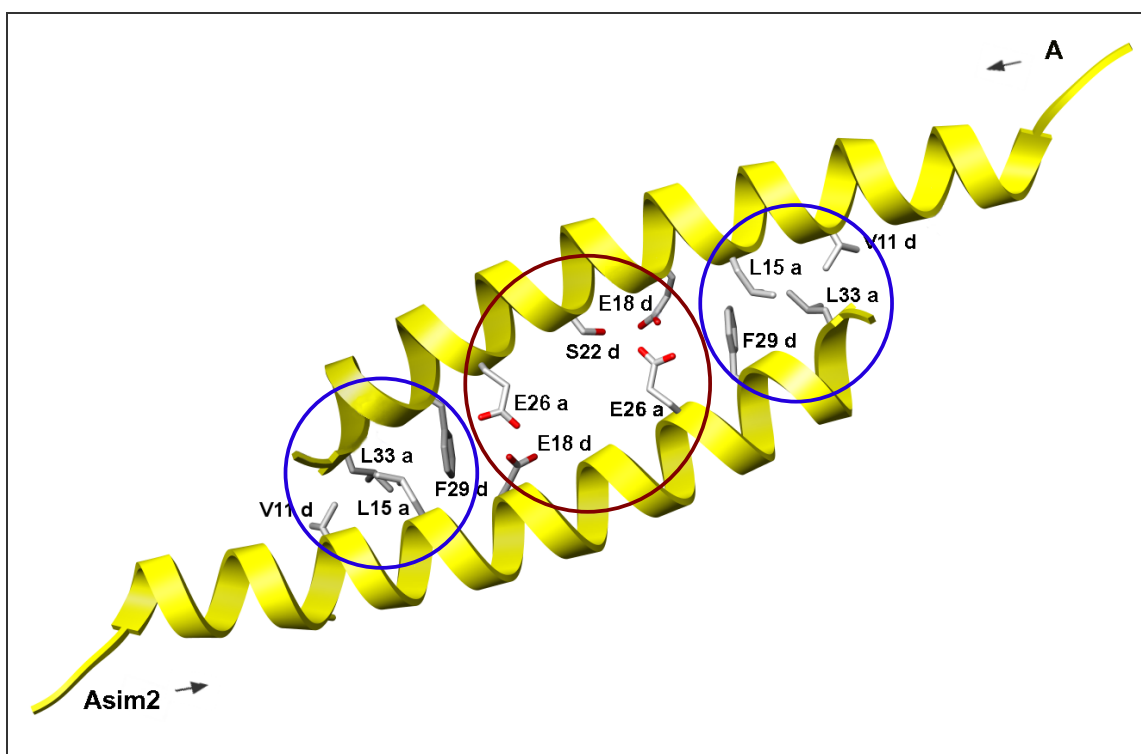


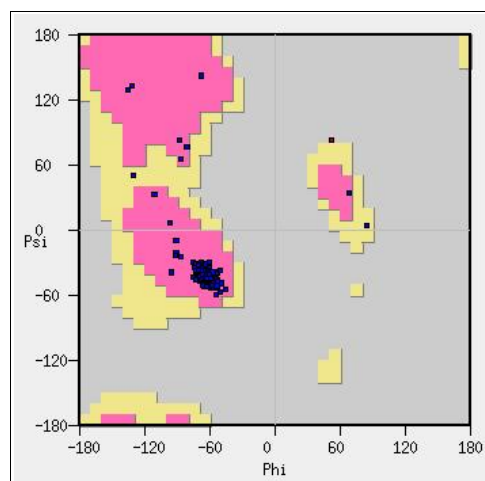
Figura 31. Estructura de las dos cadenas *A* y *Asim2*. Se muestran los contactos establecidos entre los residuos localizados en las posiciones *a* y *d*, analizados mediante el programa *Contact* de *CCP4*. Los contactos hidrofóbicos, localizados en los extremos entre los residuos V11-L15-L33-F29 se señalan en color azul, mientras que lo de carácter polar se sitúan en el centro de la doble hélice y se muestran con un círculo de color rojo. Figura realizada con el programa *Chimera*.

3.4.4 Validación de la estructura de *BRMS1*₅₁₋₈₄

La validación del ajuste de todos los parámetros que definen la estructura mediante los programas expuestos en *Materiales y Métodos*, demostró una alta fiabilidad del modelo. Los factores *R* y las desviaciones cuadráticas medias para los enlaces y los ángulos resultantes, todos encuadrados dentro de rangos razonables, se exponen en la Tabla XIII.

El diagrama de *Ramachandran*, facilitado por el programa *Coot* (Emsley y Cowtan, 2004), muestra los ángulos diedros “phi; ϕ ” y “psi; ψ ” permitidos entre los enlaces de la cadena principal $C_{i-1}-N_i-C\alpha_i-C_i$ y $N_i-C\alpha_i-C_i-N_{i+1}$ respectivamente (Kleywegt y Jones, 1996; Lovell *et al.*, 2003). En él se diferencian únicamente los 192 residuos, cuyos enlaces están caracterizados por ángulos diedros, formados por cuatro átomos, por lo que los residuos terminales de cada cadena quedan excluidos de este estudio. De ellos, 189 residuos (98,3 %) están localizados en regiones preferidas (rosas en la figura), 2 aminoácidos (1,11 %) en regiones permitidas (amarillas) y 1 aminoácido (0,56 %) (concretamente el Glutámico 9 de la cadena G) en la región energéticamente desfavorable, tal y como se puede observar en la Figura 32.

Figura 32. Diagrama de *Ramachandran* conteniendo los 192 residuos de la unidad asimétrica del cristal de BRMS1₅₁₋₈₄. En rosa se representan las regiones favorables-preferidas, en amarillo, las favorables-permitidas y en gris el área donde se localizan los residuos energéticamente desfavorable. Facilitado por el programa *Coot*.



Se realizaron validaciones en cuanto al ajuste del modelo trazado a la densidad electrónica y a la geometría permitida, usando para ello la herramienta *Validate* del programa *Coot*. Prácticamente todos los residuos modelados de la unidad asimétrica se ajustan al mapa 2 Fo - Fc. Los aminoácidos que se ajustan menos a la densidad son los situados en los extremos de todas las cadenas, a excepción de aquellos que están completamente modelados. Todos los residuos, independientemente de su localización, mantienen una geometría dentro de los parámetros permitidos (Figura 33).

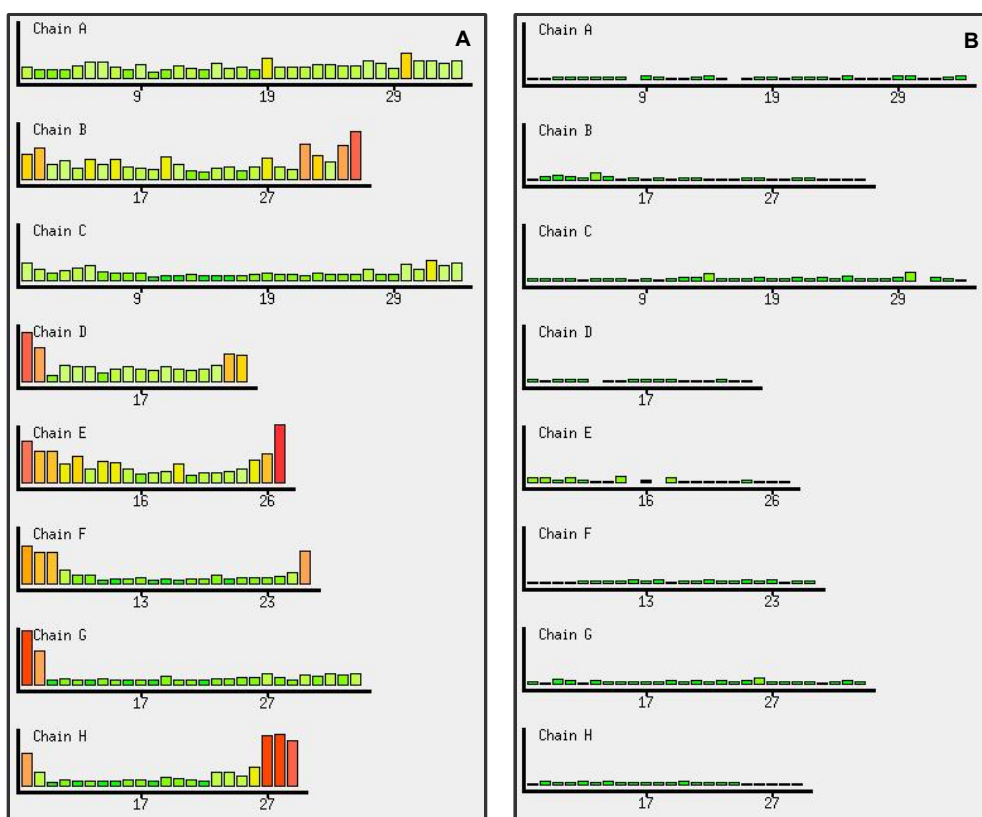


Figura 33. Validación del modelo, según el ajuste de los residuos de las ocho cadenas de la unidad asimétrica de BRMS1₅₁₋₈₄ a la densidad electrónica (A), y a la geometría (B). El nivel de ajuste se muestra según una escala de color de las barras, en la que el color verde representa mayor ajuste y progresivamente hacia el color rojo, un menor ajuste. Los gráficos se han obtenido utilizando el programa *Coot* (herramienta *Validate*).

Los residuos que presentaban menos ajuste en el mapa de densidad electrónica eran principalmente, como se ha comentado con anterioridad, los extremos de las cadenas. Estas regiones resultaron también con *factores B* o vibraciones de los átomos, elevados. En la Figura 34 se muestran en escala de colores, la variación de dichos *factores B* a lo largo de todos los átomos de las cadenas que conforman la unidad asimétrica, desde aquéllos con menor *factor B* (colores azules) a los que tienen un valor mayor (rojo) progresivamente.

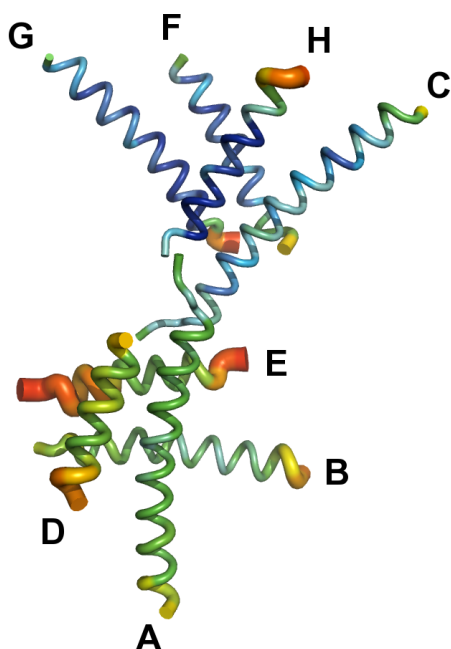


Figura 34. Representación de la variación atómica de los *factores B* de las 8 cadenas de BRMS1₅₁₋₈₄ que forman la unidad asimétrica. Se representa mediante una escala progresiva de colores, en la que el color azul se corresponde con los valores más bajos de *B* (aproximadamente 30 Å²), y los rojos con los mas elevados (unos 50 Å²).

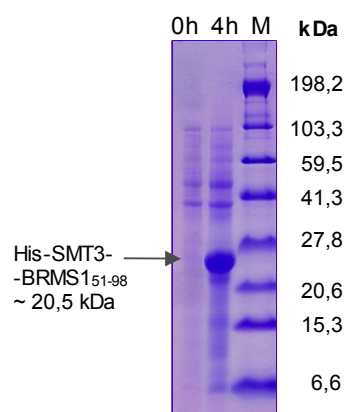
4. Expresión, purificación y cristalización de BRMS1 aa51-98.

La resolución de la estructura del motivo “coiled-coil” que acabamos de discutir en el apartado anterior, flanqueado por los aminoácidos 51 al 84, nos sugirió plantearnos la idea de intentar elongar la doble hélice antiparalela en su extremo carboxilo-terminal, con el fin de poder visualizar el giro completo de una hélice sobre la otra y así profundizar en el estudio de dicho motivo. Para ello, diseñamos una nueva construcción, concretamente de 15 residuos mas largo en el extremos carboxilo (el fragmento 51 a 98).

El ADN codificante para dicha región, fue clonado en el vector de expresión pET28-SMT3, para ser transformado y expresado en células de *E.coli* de la cepa *Rosetta(DE3)*, usando como medio de crecimiento *2xYT*. Tras añadir IPTG a una concentración final de 1 mM, se mantuvo la sobre-expresión durante 4 h. Los niveles de sobre-expresión obtenidos de la proteína de fusión His-SMT3-BRMS1₅₁₋₉₈ se muestran en la Figura 35.

Figura 35. Expresión de His-SMT3-BRMS1₅₁₋₉₈ en cepa de *E.coli*, analizada mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 12 % BIS-TRIS.

El carril marcado como “0h” se corresponde con el total de proteínas expresadas por las células antes de inducir la sobre-expresión de His-SMT3-BRMS1₅₁₋₉₈ mediante adición de IPTG. El carril “4h” contiene el total de proteínas presente en las células transcurridas 4 horas de añadir el agente inductor. Se usó un marcador (M) de alto peso molecular cuyos pesos moleculares se muestran a la derecha en unidades de kDa.



La proteína de fusión His-SMT3-BRMS1₅₁₋₉₈ fue purificada de forma similar a His-SMT3-BRMS1₅₁₋₈₄, en el apartado 3.2 de esta sección, siguiendo el protocolo expuesto en *Materiales y Métodos* (sección 3.1), a través de cromatografía de afinidad a histidinas, donde se eluyó utilizando un gradiente lineal de Imidazol. A continuación, las fracciones que contenían la proteína de interés, se sometieron a lisis con la proteasa *ULP* tras la que se pasó por una segunda columna de afinidad a histidinas con el fin de eliminar los productos originados de la digestión, ajenos a la secuencia de interés (datos no mostrados). Finalmente, la fracción no retenida de esta columna se inyectó en una columna de exclusión por tamaño molecular del tipo “Superdex 75 26/60”, donde el volumen de elución máximo fue de aproximadamente 160 mL en un tampón con 20 mM TRIS pH 7,5 y 50 mM NaCl. Tras analizar las fracciones en un gel de poliacrilamida SDS-PAGE teñido con azul de Coomassie, observamos que dicho máximo correspondía a BRMS1₅₁₋₉₈, como se muestra en la Figura 36, aunque también fue identificado mediante análisis de proteómica.

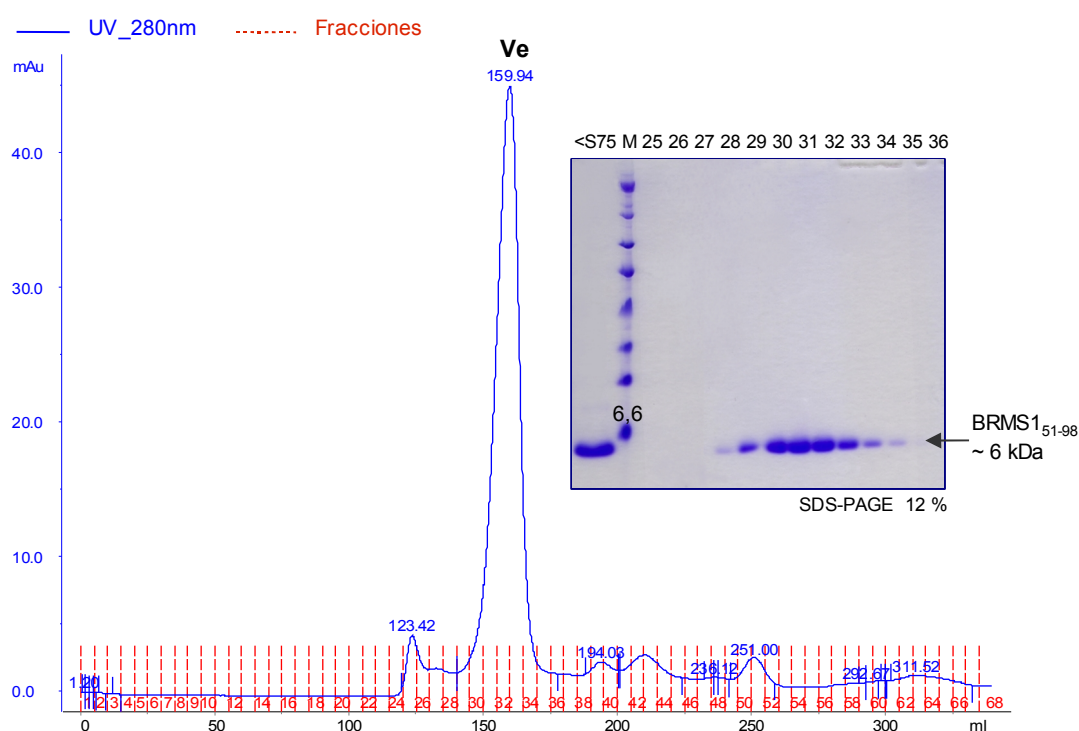
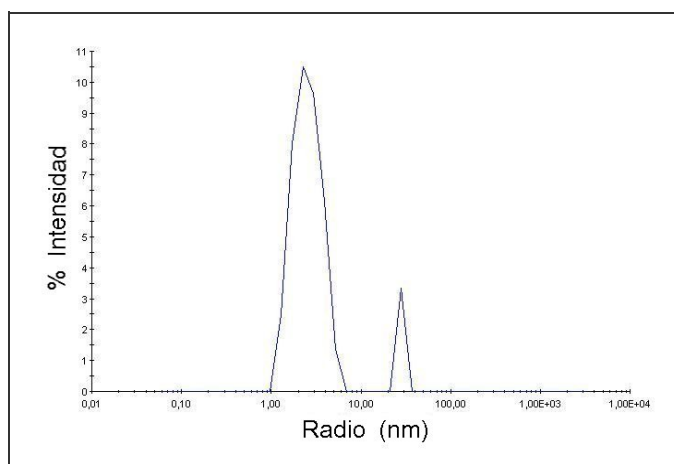


Figura 36. Cromatografía de exclusión por tamaño molecular de BRMS1₅₁₋₉₈ libre de His-SMT3. El cromatograma muestra la absorbancia a una longitud de onda de 280 nm (señal azul), de las proteínas eluidas en distintos volúmenes según su tamaño y recogidas en fracciones de 5 mL representadas en rojo. Se muestra sobre cada pico de elución, su correspondiente volumen, siendo “Ve” el de la proteína de interés. Las muestras antes de la inyección en la columna (<S75) y eluidas en el pico mayoritario (28-35) se muestran en el gel de poliácridamida (SDS-PAGE) BIS-TRIS del 12 % (*Invitrogen*^l). La electroforesis fue llevada a cabo en presencia de tampón MES. M: marcador, donde se representa el tamaño en kDa del peso más cercano a la banda de interés.

La interpolación en una curva de calibración realizada con estándares de tamaños conocidos, del volumen correspondiente al máximo de absorbancia a 280 nm (“Ve” en la figura), sugirió un peso molecular aproximado de 32 kDa. Este tamaño molecular se explicaría mediante la presencia de una conformación pentamérica, ya que el peso de cada monómero son unos 6 kDa.

Estos resultados se confirmaron mediante ensayos de dispersión dinámica de la luz (DLS), en los que utilizando el polipéptido a una concentración de 2 mg/mL en el tampón de la columna de filtración en gel (Tabla III-E), se estimó un tamaño molecular de 33 kDa, con un radio hidrodinámico de 2,6 nm y una polidispersidad de la muestra de 35 % (Figura 37).

Figura 37. Dispersión Dinámica de la Luz de BRMS1₅₁₋₉₈, a una concentración de 2 mg/mL, en un tampón conteniendo 20 mM TRIS pH 7,5 y 50 mM NaCl. Se muestra un máximo de intensidad correspondiente a una población con una polidispersidad del 35 % y un tamaño molecular de 33 kDa.



Las fracciones 29 a 34 procedentes de la filtración en gel fueron agrupadas y concentradas mediante unidades de filtración de 5 kDa de diámetro de poro, hasta alcanzar una concentración de 105 mg/mL, en el tampón conteniendo 20 mM TRIS pH 7,5 y 50 mM NaCl. El estado de pureza de la proteína final usada para los ensayos de cristalización, cuya identidad fue comprobada mediante espectrometría de masas, se muestra en la Figura 38.

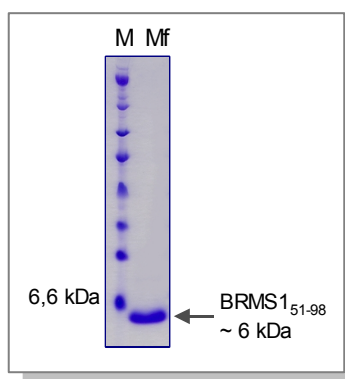


Figura 38. Gel de poliacrilamida al 12 %, BIS-TRIS, mostrando el elevado grado de pureza de la proteína final BRMS1₅₁₋₉₈ (muestra final; Mf), utilizada para ensayos de cristalización, a una concentración de 150 mg/mL y en un tampón de 20 mM TRIS pH 7,5 y 50 mM NaCl. Se usó un marcador (M) de estándares de alto peso molecular (BioRad^E). Se muestran los tamaños moleculares en kDa de las bandas de interés.

Para realizar los ensayos de cristalización mediante difusión de vapor, se mezclaron determinados volúmenes de la proteína BRMS1₅₁₋₉₈ a una concentración de 105 mg/mL, con volúmenes iguales de unas 800 condiciones diferentes de cristalización. Al cabo de 3 días de incubación a una temperatura de 22 °C, se obtuvieron cristales con forma cúbica, como los que se exponen en la Figura 39-A. La gota donde aparecieron los cristales, contenía la condición comercial 0,1 M citrato sódico pH 5,5 y 2 M sulfato amónico, y un volumen final de 0,2 µL. Estos cristales se optimizaron disminuyendo la concentración de proteína a 50 mg/mL, variando la condición de cristalización a 0,1 M citrato sódico pH 5,25 y 1,5 M sulfato amónico y aumentando el volumen final de gota a 2 µL. Con ello se consiguieron cristales del tipo expuesto en la Figura 39-B, con unas dimensiones aproximadas de 0,25 x 0,25 x 0,25 mm.

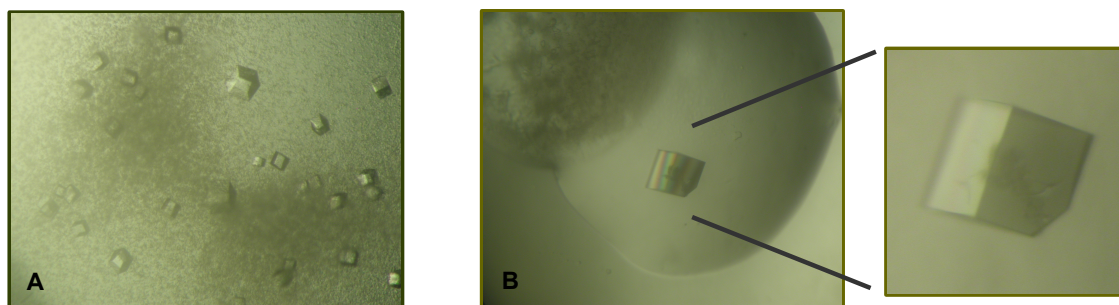


Figura 39. Cristales de BRMS1₅₁₋₉₈ obtenidos a una temperatura de 22 °C, mediante la técnica de difusión de vapor en:

A) “gota sentada” en formato de 96 pocillos, en ensayos iniciales de cristalización, y mezclando 0,1 μ L de proteína con 0,1 μ L de reservorio, en 0,1 M citrato sódico pH 5,5 y 2 M sulfato amónico.

B) “gota colgante” usando formato de 24 pocillos, tras mezclar volúmenes iguales (1 μ L) de proteína y reservorio, en 0,1 M citrato sódico pH 5,25 y 1,5 M sulfato amónico. Este cristal optimizado fue usado para ensayos de difracción.

Dicho cristal (Figura 39-B) se sometió a difracción de rayos X de la línea BM16 del sincrotrón de Grenoble, tras ser bañado con Paratone-N y enfriado rápidamente con gas de nitrógeno a 130 K. Se recogió un juego completo de datos de difracción a una resolución máxima de 2,26 Å (Figura 40). Para ello se utilizaron los parámetros que se exponen a continuación en la Tabla XV.

Energía	Distancia	Resolución	Incremento del ángulo ϕ total	Incremento del ángulo ϕ / Imagen	Tiempo de exposición
1,0071 Å (12,3110 keV)	217,4 mm	48,34 - 2,26 Å	56 ° 100 °	0,5 °	5 seg.

Tabla XV. Parámetros usados en la línea BM16 del sincrotrón, para la recolección del juego de datos de difracción de un cristal de BRMS1₅₁₋₉₈. Se muestran por orden, la longitud de onda y energía con la que se emite el haz de rayos X incidente sobre el cristal; la distancia entre el cristal y el detector y los límites de resolución a la que se recogieron las difracciones emitidas por las moléculas del cristal. A continuación se muestran parámetros del ángulo “phi” (ϕ) que determina el posicionamiento del goniómetro donde se soporta el cristal. Estos son el ángulo total colectado y el incremento en la rotación del mismo en cada imagen. También se expone el tiempo de exposición en segundos para la toma de cada imagen.

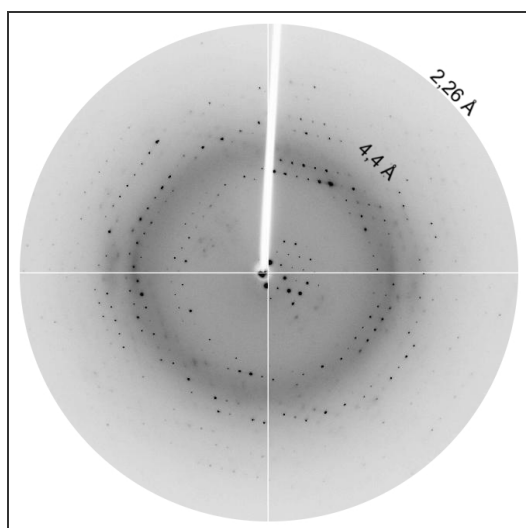


Figura 40. Imagen representativa, correspondiente a un ángulo de 0,5 ° del patrón de difracción obtenido en la recolección de los datos del cristal de BRMS1₅₁₋₉₈. El cristal fue embebido en Paratone-N antes de proceder al enfriamiento a 130 K para difractarlo con rayos X de una longitud de onda de 1 Å y una distancia de 217 mm. Se muestran sobre la imagen algunos límites de resolución.

Tras la reducción de los datos obtenidos, se obtuvieron los valores estadísticos expuestos en la Tabla cristalográfica XVI.

BRMS1 aa (51-98)	
Reducción de datos	
Grupo espacial	<i>H</i> 32
Dimensiones de la celda	
a; b; c (Å)	60,039; 60,039; 131,220
α, β, γ (°)	90; 90; 120
Rango de resolución (Å) (última capa)	48,34 - 2,26 (2,38 - 2,26)
Observaciones (únicas)	39699 (4445)
Complejidad (%) (última capa)	99,2 (100)
Multiplicidad (última capa)	8,9 (8,9)
R _{sym} ^a (última capa)	0,105 (0,186)
R _{meas} ^b (última capa)	0,112 (0,197)
R _{pim} ^c (última capa)	0,037 (0,065)
<I/σ> (última capa)	17,5 (10,3)

^a R_{sym} es el *factor R* no sopesado sobre las intensidades simétricamente relacionadas.

$R_{sym} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, donde $I_i(hkl)$ es la intensidad de la reflexión hkl y $\langle I(hkl) \rangle$ es la intensidad media de todas las observaciones i de la reflexión hkl .

^b $R_{meas} = \{ \sum_{hkl} [N/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| \} / \sum_{hkl} \sum_i I_i(hkl)$, donde $I_i(hkl)$ son las intensidades observadas, $\langle I(hkl) \rangle$ son las intensidades medias y N es la multiplicidad de la reflexión hkl .

^c $R_{pim} = \{ \sum_{hkl} [1/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| \} / \sum_{hkl} \sum_i I_i(hkl)$, donde $I_i(hkl)$ son las intensidades observadas, $\langle I(hkl) \rangle$ son las intensidades medias y N es la multiplicidad de la reflexión hkl .

<I/σ> es la media de la relación entre la intensidad de la difracción y el ruido de fondo.

Tabla XVI. Resultados estadísticos cristalográficos obtenidos tras el procesamiento y simplificación de los datos de difracción recogidos del cristal de BRMS1₅₁₋₉₈.

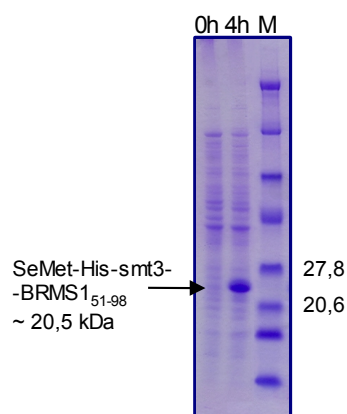
Con el fin de obtener las fases que nos permitiera resolver la estructura de BRMS1₅₁₋₉₈, se abordaron diferentes técnicas. Se intentó, realizar reemplazo molecular, usando como modelo las posiciones atómicas de la estructura ya resuelta de BRMS1₅₁₋₈₄. Para ello se usaron las dos cadenas antiparalelas *A-A_{sim2}*, completas o acotando sólo a la superficie de contacto entre las dos cadenas. También se usaron diferentes tamaños de una de las cadenas de forma individual, y utilizando distintos porcentajes de similitud en la entrada de *Refmac5*. De igual manera se realizaron numerosos reemplazos con moléculas estructuradas en “coiled-coil”, localizadas en la base de datos *PDB* (“*Protein Data Bank*”), o bien, con cierta homología con la secuencia de BRMS1₅₁₋₉₈, de la misma base de datos. Desafortunadamente, ninguna de estas aproximaciones resolvieron en problema de resolver las fases.

En el mismo sentido, se intentó conseguir cristales derivados de mercurio, platino y oro, mediante impregnación del cristal en dichas soluciones. Se obtuvieron escasos cristales estables y no dañados, pero tras obtener los datos de difracción, la señal anómala y diferencias isomorfas del derivado no fueron suficientes para obtener las fases.

Por todo ello, finalmente se procedió a la sobre-expresión de dicho polipéptido BRMS1₅₁₋₉₈ con residuos de Selenometionina (SeMet) la cual, añadida en el medio de crecimiento de las células durante la expresión, es incorporada a la secuencia polipeptídica, generando señal anómala (Se) que puede ser usada para resolver las fases. La secuencia aminoacídica de esta construcción contiene una metionina en la posición 14, que es sustituida por metionina con un átomo de selenio en lugar de azufre. Para su expresión, el vector pET28-SMT3-BRMS1₅₁₋₈₄ se transformó en células de la cepa *Rosetta(DE3)*, con el fin de expresar SeMet-His-SMT3-BRMS1₅₁₋₉₈, usando un medio mínimo complementado con SeMet (protocolo en la sección de *Materiales y Métodos*). Finalmente, tras cuatro horas de inducción con 1 mM de IPTG a 37 °C, el cultivo celular obtenido mostró unos niveles de sobre-expresión aceptables para el abordaje de ensayos de cristalización. Dichos niveles se muestran en la Figura 41.

Figura 41. Expresión de His-SMT3-BRMS1₅₁₋₉₈ con selenometionina, en cepa de *E.coli* (*Rosetta(DE3)*), utilizando un protocolo para cepas no auxótrofas. La sobre-expresión se visualiza en un gel de poliacrilamida (SDS-PAGE) del 12 % BIS-TRIS.

La muestra recogida antes de inducir la sobre-expresión es el carril “0h”, mientras que el carril “4h” contiene el total de proteínas presente en las células transcurridas 4 horas de añadir el agente inductor. Se usó un marcador (M) de alto peso molecular, que se muestran a la derecha en unidades de kDa.



Los procesos de purificación y obtención de cristales, se realizaron siguiendo el mismo protocolo que para la proteína sin derivar, como ya hemos descrito al principio de este mismo apartado. Finalmente, se consiguieron cristales, con una morfología y calidad similares a los de la proteína nativa, que fueron irradiados con un haz de rayos X de la línea BM16 del sincrotrón de Grenoble, (parámetros usados no mostrados). Desafortunadamente, se obtuvo un patrón de difracción muy pobre y difuso (Figura 42), por lo que se hizo imposible procesar estos datos y con ello resolver la estructura.

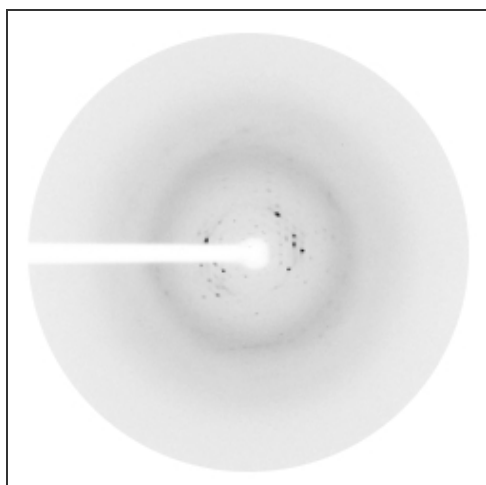


Figura 42. Ejemplo de difracción de un cristal de SeMet-BRMS1₅₁₋₉₈. Imagen recogida en la línea BM16 del sincrotrón de Grenoble. La distancia entre el cristal y el detector es de 217,4 mm, con el fin de recoger una resolución de hasta 2,26 Å. La imagen corresponde a una rotación del ángulo ϕ de $^{\circ}$.

5. *Expresión y purificación de NMI aa16-64, proteína de interacción con BRMS1.*

Una vez resuelta la estructura del motivo “coiled-coil” de BRMS1, nos centramos en la obtención del mismo en complejo con otras proteínas.

Para este estudio, seleccionamos la proteína NMI, que previamente había sido rescatada en el laboratorio, como proteína que interacciona con BRMS1₅₁₋₈₄, por el Dr. José Rivera, mediante ensayos de doble híbrido. NMI es un co-factor de la transcripción que une moléculas de relevancia biológica como son N-myc o c-Myc, entre otros factores transcripcionales. La región mínima de interacción de NMI con BRMS1 fue la comprendida entre los aminoácidos 1 a 104, que contiene un motivo predicho como “coiled-coil” (del aminoácido 21 al 64), y un dominio IFP35 (residuos 29 al 104) y una segunda región comprendiendo los dos dominios NID, representados en la Figura 10. La presencia de este motivo “coiled-coil” en el primer fragmento rescatado, sugirió la posibilidad de que el reconocimiento entre ambas proteínas estuviese mediado vía coiled-coil, por lo que abordamos el clonaje y expresión de NMI₁₆₋₆₄ para caracterizar la interacción y realizar ensayos de cristalización.

Para abordar el clonaje del motivo “coiled-coil” de NMI, diseñamos diferentes construcciones en torno a esta región. Cuatro de estas construcciones fueron diseñadas elongando y acortando 5 residuos en ambos extremos terminales, con el fin de evitar una acotación equivocada de la predicción del motivo. Uno de estos fragmentos fue el comprendido entre los aminoácidos 16 y 64, que fue el único que se consiguió expresar y se obtuvo soluble, por lo que nos centramos en él.

5.1 *Expresión*

Para expresar el motivo “coiled-coil”, se transformó el vector de expresión pET28-SMT3, conteniendo la secuencia que codifica para el fragmento 16 al 64 de NMI, en

células de *E.coli* de la cepa *BL21(DE3)*. Las condiciones óptimas de sobre-expresión de la proteína His-SMT3-NMI₁₆₋₆₄ en este sistema procariota, fueron usando un medio de crecimiento *2xYT*, a 37 °C e induciendo con 1 mM de IPTG. En la Figura 43 se compara visualmente el nivel de proteína sobre-expresada, en un gel de poliacrilamida (SDS-PAGE) tras 4 horas de añadir el agente inductor.

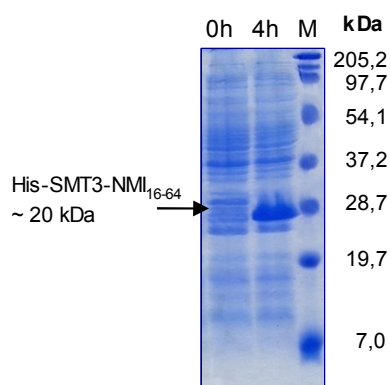


Figura 43. Expresión de His-SMT3-NMI₁₆₋₆₄ en cepa de *E.coli*, analizada mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 17 % TRIS-HCl.

El carril marcado como “0h” se corresponde con el total de proteínas expresadas por las células antes de inducir la sobre-expresión de NMI. Transcurridas 4h de añadir el agente inductor se alcanzaron los niveles de His-SMT3-NMI₁₆₋₆₄ expuestos en el carril “4h”, cuya banda se muestra señalada, con un tamaño aproximado de 20 kDa. A la derecha se muestra el marcador usado (M).

5.2 Purificación

El cultivo celular obtenido tras la sobre-expresión de His-SMT3-NMI₁₆₋₆₄ se sonicó para su lisis en un tampón que contenía 50 mM TRIS pH 7,5; 0,5 M NaCl y 0,5 % (v/v) Tween-20. Posteriormente se separó la fracción soluble de la insoluble (“S” e “I” respectivamente en la Figura 44) mediante ultracentrifugación. La fracción soluble contenía la proteína de fusión de interés, como se puede ver en la misma figura, que fue utilizada para realizar una primera purificación a través de una columna de afinidad por histidinas.

Tras inyectar la fracción soluble conteniendo la proteína His-SMT3-NMI₁₆₋₆₄ en una columna de afinidad a histidinas, se aplicó un gradiente lineal de Imidazol, hasta 0,5 M, en el tampón de lavado, que contiene 50 mM TRIS pH 7,5 y 150 mM NaCl, y se recogieron fracciones de 4 mL de volumen (“A4-A12” en la Figura 44-B). La proteína de interés eluyó en dos máximos de absorción a 280 nm, que comprende concentraciones de imidazol de entre 156 mM y 211 mM (31 y 42 % tampón B o de elución). Estos resultados, así como las fracciones obtenidas en cada etapa, se muestran en la Figura 44.

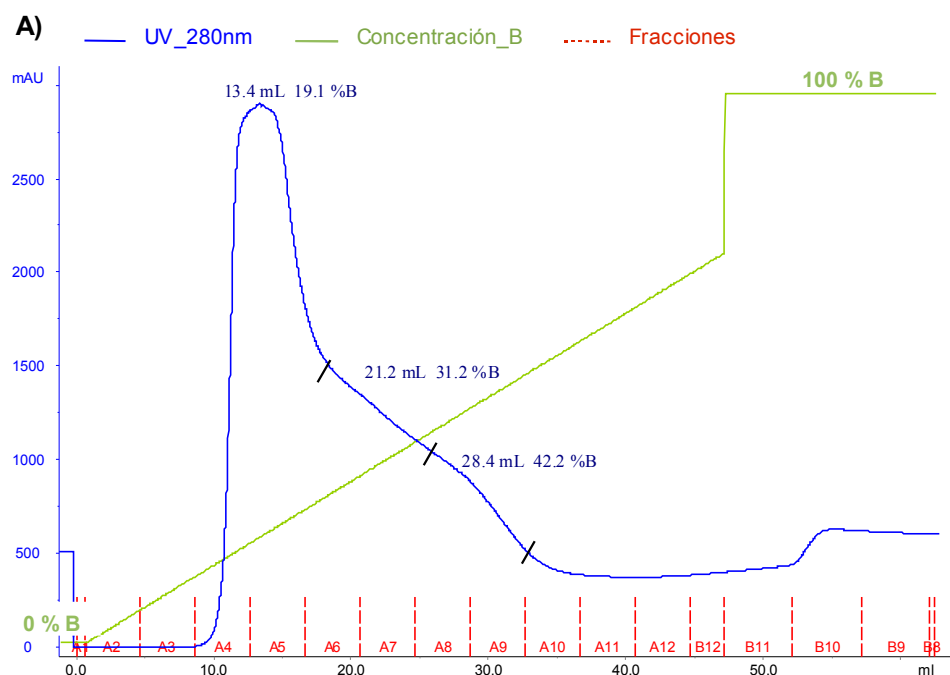
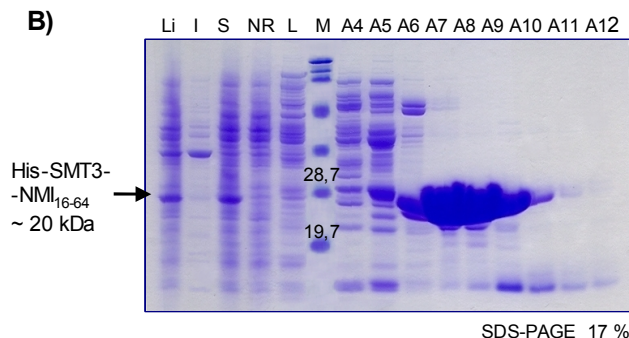


Figura 44. Cromatografía de afinidad de His-SMT3-NMI₁₆₋₆₄. **A)** Se muestra en el cromatograma el patrón de absorción de luz ultravioleta a una longitud de onda de 280 nm, obtenido tras la elución de las proteínas unidas a la columna de afinidad. Sobre el máximo de cada pico se exponen el volumen de elución y el porcentaje de tampón de elución (B) en los mismos. Dos rangos diferentes de elución de NMI (31,2 y 42,2 % B) se separan por líneas diagonales.

B)



B) Electroforesis en gel de poliacrilamida al 17 % en TRIS-HCl, de las muestras obtenidas en la cromatografía. Antes del carril cargado con el marcador de alto peso molecular (M) se muestran por orden y de izquierda a derecha, la muestra inicial de células expresando NMI₁₆₋₆₄ lisadas (Li), las fracciones insoluble y soluble (“I” y “S” respectivamente) tras la lisis y ultracentrifugación, la fracción no retenida en la columna (NR) y la expulsada de la columna en el lavado (L). Tras el marcador se localizan las fracciones eluidas con un gradiente continuo de Imidazol (A4-A12). La banda correspondiente a la proteína de interés se muestra con una flecha.

Las muestras correspondientes a las fracciones A6-A9 fueron englobadas en un único volumen de 20 mL con una concentración de 5 mg/mL, para proceder a eliminar la cola de histidinas por proteólisis con la proteasa *ULP*. Para ello se le añadió a la muestra antes de la proteólisis (“<C” en la Figura 45), 100 µg de proteasa *ULP*. Este proceso se llevó a cabo dializando contra el tampón con 50 mM TRIS pH 7,5 y 0,2 M NaCl, a 4 °C de temperatura, durante 14-16 horas.

De esta digestión (representada como “C” en la Figura 45), resultaron dos productos bien diferenciados, correspondientes a la proteína SMT3 unida a histidinas (His-SMT3), con un tamaño aproximado de 14 kDa, y al fragmento de interés de NMI libre de fusión (NMI₁₆₋₆₄), de unos 5,8 kDa de peso molecular. Como en los casos

anteriores, la eficacia del corte mediado por la proteasa *ULP* fue muy elevada, digiriéndose prácticamente la totalidad de la proteína de fusión.

Con el fin de retener tanto la proteína de fusión que no ha sido digerida (His-SMT3-NMI₁₆₋₆₄), como el fragmento originado (His-SMT3) y la enzima usada para la proteólisis que presenta una fusión a histidinas, la muestra procedente de la proteólisis se pasó por una segunda columna de afinidad por histidinas cargada con níquel. Como era de esperar, el fragmento no retenido en la columna de afinidad (“NR” en la Figura 45), se corresponde con el polipéptido NMI₁₆₋₆₄ libre de His-SMT3, como se puede comprobar tras análisis mediante SDS-PAGE.

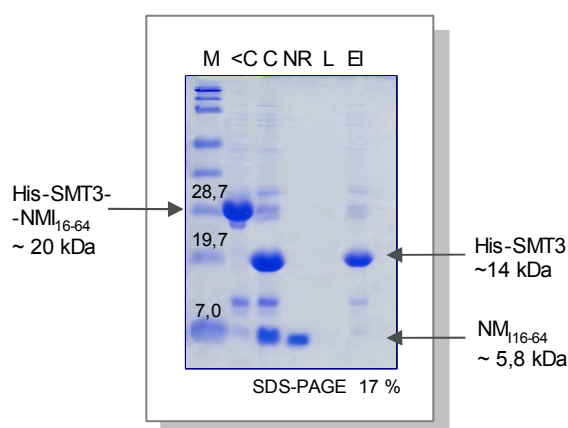


Figura 45. Muestras resultantes de la proteólisis de His-SMT3-NMI₁₆₋₆₄ analizadas mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 17% - TRIS-HCl.

Los carriles “<C” y “C” se corresponden con muestras tomadas de la mezcla proteica antes y después de la proteólisis respectivamente. “NR” hace mención a la muestra no retenida en la columna de afinidad (NMI₁₆₋₆₄), y “L” al lavado realizado antes del paso de elución (E).

“M” corresponde al marcador, tres de los cuales se muestran los pesos en kDa.

His-SMT3-NMI₁₆₋₆₄ es la proteína fusionada a histidinas y His-SMT3 es el fragmento de fusión a NMI₁₆₋₆₄, los cuales han sido separados tras la proteólisis. Se muestran los tamaños de cada proteína en números aproximados (~).

Esta fracción no retenida en la columna (de un volumen aproximado de 22 mL a una concentración de 1 mg/mL), fue concentrada hasta 4,5 mL para ser cargada en un lazo de 5 mL de capacidad de una columna de filtración en gel del tipo “Superdex 75 26/60”, previamente equilibrada con un tampón de 20 mM TRIS pH 7,5 y 50 mM NaCl.

La cromatografía se llevó a cabo en el citado tampón, y las proteínas eluidas se recogieron en fracciones de 5 mL de volumen. Debido a la ausencia de aminoácidos aromáticos en el polipéptido NMI₁₆₋₆₄, que absorben la luz ultravioleta a una longitud de onda de 280 nm, se dificultó el seguimiento del mismo, en lo que respecta a su visualización e identificación en un cromatograma y a su cuantificación usando esta longitud de onda. Por ello, se hizo necesario analizar todas las fracciones de la cromatografía de exclusión por tamaño molecular, mediante electroforesis en gel de poliacrilamida, aunque sólo se muestren aquéllas conteniendo la banda de interés en la Figura 46. A partir de la proteólisis, NMI₁₆₋₆₄ se cuantificó usando el método colorimétrico “Bradford”.

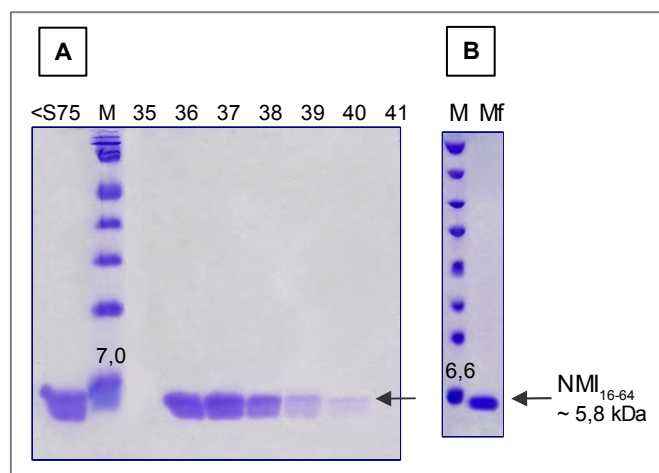


Figura 46. **A)** Filtración en gel de NMI₁₆₋₆₄ tras la eliminación de la cola de histidinas. Se muestra el gel de poliacrilamida (SDS-PAGE) TRIS-HCl del 17 % con las fracciones 36 a 40 conteniendo NMI₁₆₋₆₄. La muestra antes de la inyección en la columna (<S75) se expone a la izquierda del marcador (M). **B)** Gel de poliacrilamida del 4-12 %, BIS-TRIS en tampón MES, mostrando un elevado grado de pureza de la proteína final NMI₁₆₋₆₄ (muestra final; “Mf”) en un tampón conteniendo 20 mM TRIS pH 7,5 y 50 mM NaCl, utilizada para los ensayos de cristalización. Se utilizó en ambos casos un marcador (M) de estándares de alto peso molecular.

Están representados en kDa sobre su correspondiente banda, solamente el peso del estándar cercano a la banda de interés. La banda marcada con una flecha, cuyo peso molecular es de aproximadamente 5,8 kDa, corresponde a la proteína NMI₁₆₋₆₄.

El volumen de elución medio (185 mL) correspondiente a las fracciones 36 a 40, fue usado para calcular el tamaño molecular aproximado del polipéptido de interés. De esta forma, tras interpolar dicho valor en una curva patrón con estándares conocidos, sugiere que su conformación oligomérica más probable, resultando un tamaño aproximado de 17,5 kDa, sería trimérica.

Las fracciones 36 a 40 fueron agrupadas en un volumen de 25 mL y una concentración aproximada de 0,3 mg/mL, en tampón 20 mM TRIS pH 7,5 y 50 mM NaCl. Esta proteína final purificada, cuyo elevado grado de pureza se muestra en el gel B de la Figura 46, fue usada para diferentes experimentos, para los que se trató de forma distinta. Para estudios de interacción proteína-proteína con BRMS1₅₁₋₈₄, explicados en el siguiente apartado, la muestra NMI₁₆₋₆₄ se llevó a concentraciones de 300 μ M y 95 μ M, para experimentos de RMN (espectro bidimensional) e ITC respectivamente.

También se realizaron ensayos de cristalización, tanto de NMI₁₆₋₆₄ individual, como mediante co-cristalización con BRMS1₅₁₋₈₄, usando concentraciones equimolares con dicha proteína.

Con el fin de caracterizar biofísicamente el comportamiento de NMI₁₆₋₆₄ en solución, se realizaron ensayos de Dicroísmo Circular. Para ello se recogió a 20 °C de temperatura, un espectro de absorción en un rango de longitud de onda de 190 nm a 250 nm sobre la proteína NMI₁₆₋₆₄ a una concentración de 0,5 mg/mL. La curva obtenida (Figura 47) mostró dos mínimos, uno mas acentuado en aproximadamente 210 nm y el segundo en 222 nm. Tras comparar este comportamiento con los característicos de las estructuras secundarias, observamos una conformación mayoritaria helicoidal, ya que adquiere se aproxima a los dos mínimos característicos de dicha conformación.

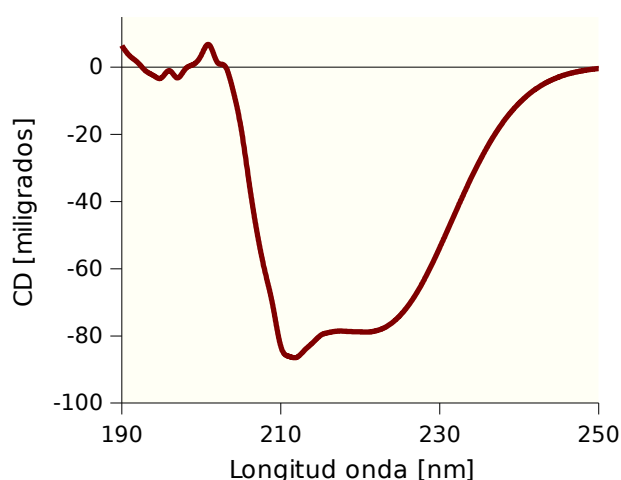


Figura 47. Espectroscopia de Dicroísmo Circular del polipéptido NMI₁₆₋₆₄.

Se muestra el promedio en elipticidad (o unidad de CD en miligrados) obtenido con una concentración de 0,5 mg/mL, en un rango de longitud de onda de 190 nm a 250 nm. La muestra fue analizada en un tampón con 10 mM HEPES pH 7,5 y 50 mM NaCl, a 20 °C de temperatura. La curva de absorción muestra dos mínimos localizados en longitudes típicas de las estructura α -hélice, uno en 208 nm y el segundo en aproximadamente 222 nm.

6. Caracterización de las interacciones proteína-proteína: Evidencias en la formación de complejos.

La formación de complejos y con ello la caracterización de ciertas interacciones, expuestas a continuación, se convirtió en un objetivo indirecto de este trabajo, con una única finalidad, llevar estos complejos a ensayos cristalográficos. Por este motivo, pese a previamente tener evidencias de la interacción entre las proteínas de interés, todas las interacciones entre las construcciones proteicas utilizadas en este estudio, intentaron ser comprobadas experimentalmente en el laboratorio mediante diversos ensayos “*in vitro*”.

El poder caracterizar las interacciones que nos encuadran y, en la medida de lo posible, poder ofrecer valores estequiométricos, representarían un importante logro para el diseño de futuros experimentos y estudios cristalográficos.

A pesar de haber intentado la formación de complejos entre todas aquellas construcciones clonadas que fueron expresadas, solubles y no agregadas, solamente se expondrán en este apartado las interacciones en las que nos centramos, por obtener evidencias en los ensayos que se explican a continuación, que apuntan a las formaciones de estos complejos. Con ello, principalmente se estudiaron e intentaron obtener los complejos formados por BRMS1_{FL} + NMI_{FL} y BRMS1₅₁₋₈₄ + NMI₁₆₋₆₄.

6.1 Co-expresión y co-purificación del complejo BRMS1 FL + NMI FL (secuencias completas)

En primer lugar nos centramos en el estudio, con fines cristalográficos, de las secuencias completas de las proteínas humanas BRMS1 y NMI. Para tal fin, pretendimos obtener evidencias de la interacción de las construcciones que posteriormente iban a ser utilizadas para ensayos de cristalografía, a través de alguna de las etapas llevadas a cabo en la expresión y/o purificación.

Por este motivo, las secuencias codificantes de ADN para las proteínas completas BRMS1 y NMI, fueron clonados en los vectores de expresión pET28-SMT3 y pET21

respectivamente. En el caso de NMI, se introdujo un codón de parada entre la secuencia proteica y la cola de histidinas, con el fin de expresar la proteína sin unir a histidinas, mientras que la proteína BRMS1 se pretendía expresar unida a His-SMT3. Ambos plásmidos fueron co-transformados en células *Rosetta(DE3)* para ser expresados de forma simultánea. Tras inducir la sobre-expresión con 1 mM de IPTG, durante 5 horas y a 37 °C de temperatura, el cultivo celular fue lisado mediante sonicación, utilizando el tampón 50 mM TRIS pH 7,5; 0,5 M NaCl y 0,5 % (v/v) Tween-20, y posteriormente centrifugado a alta velocidad para separar la fracción soluble de la insoluble. Los análisis realizados sobre dichas fracciones mediante SDS-PAGE, mostraron suficientes niveles de expresión de ambas proteínas en la fracción soluble (Figura 48), para llevar a cabo las siguientes etapas de purificación y obtener elevados niveles de proteínas para ensayos de cristalización.

La fracción soluble se cargó en una columna de afinidad por histidinas, y tras ser lavada con un tampón 50 mM TRIS pH 7,5; 0,15 M NaCl y 10 mM Imidazol, se procedió a la elución con un gradiente lineal hasta 0,5 M de Imidazol (de 0 a 100 % de tampón de elución B). Tanto la proteína His-SMT3-BRMS1_{FL} como NMIFL quedaron retenidas en la columna de afinidad, eluyendo en las mismas fracciones, a una concentración de 220 mM de Imidazol (Figura 48). Las bandas correspondientes a las dos proteínas fueron aisladas del gel para ser verificadas por análisis de proteómica.

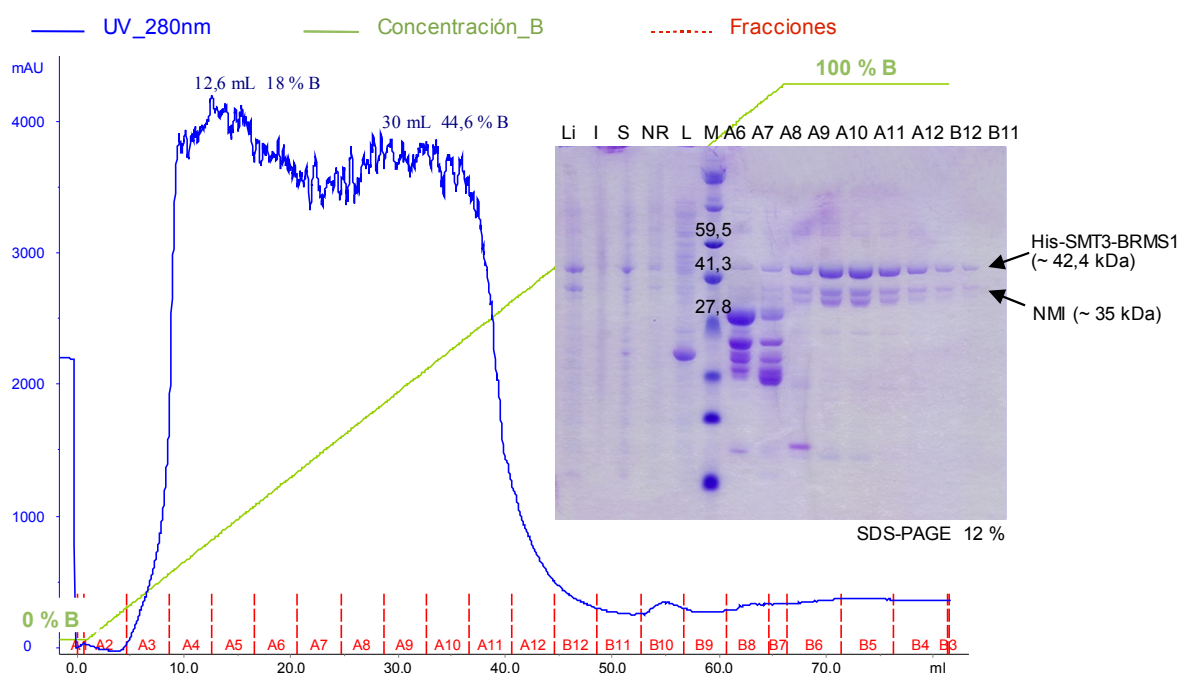
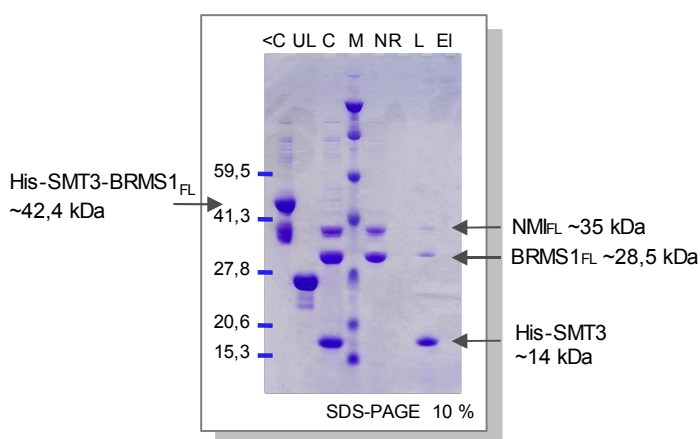


Figura 48. Cromatografía de afinidad por histidinas de las proteínas completas BRMS1 y NMI co-expresadas. La muestra resultante de la lisis celular (Li) fue centrifugada obteniéndose las fracciones insoluble y soluble (“I” y “S”, respectivamente). La muestra soluble fue inyectada en la columna cargada con níquel, que fue lavada posteriormente con el tampón de la Tabla III-B (“L”; fracción expulsada en el lavado). “NR” hace referencia a la fracción no retenida en la columna. Finalmente, ambas proteínas eluyeron a un porcentaje de 45 % (fracciones A8-B11) de tampón B (tampón de elución), correspondiente al segundo máximo del cromatograma que muestra la absorbancia a 280 nm, tras aplicar un gradiente lineal hasta 0,5 M de Imidazol. Todas estas muestras fueron visualizadas a través de electroforesis en gel BIS-TRIS al 12 % de poliacrilamida (*Invitrogen*^l), de donde se rescataron las bandas de ambas proteínas, señaladas con una flecha, para ser identificadas mediante ensayos de proteómica. Se representan, en unidades de kDa, algunos de los marcadores usados, de alto peso molecular (M).

El que NMI_{FL}, que carece de cola de histidinas, quede retenida en la columna a través de His-SMT3-BRMS1_{FL}, parece indicar que existe interacción entre ambas proteínas. Además, se comprobó el comportamiento, tanto de His-SMT3-BRMS1_{FL}, como de NMI_{FL} cargadas de forma individual en la misma columna y, si bien, His-SMT3-BRMS1_{FL} presentaba el mismo comportamiento, eluyendo en las mismas fracciones, NMI_{FL} no se unía a la columna (datos no mostrados),

Por otro lado, tras la eliminación de la cola de histidinas de la construcción de BRMS1 con la proteasa *ULP* y posterior cromatografía de afinidad con el extracto proteico, ninguna de las proteínas quedó retenida en la columna, y además eluyeron en la misma fracción (Figura 49). Estos datos refuerzan la hipótesis de que ambas proteínas están interaccionando bajo las condiciones utilizadas en esta etapa de la purificación.

Figura 49. Proteólisis de la mezcla proteica His-SMT3-BRMS1_{FL}+NMI_{FL} presente en las fracciones A8-B11 de la primera cromatografía de afinidad, con la proteasa *ULP* y purificación de la muestra digerida a través de una segunda columna de afinidad a histidinas. Se muestra un gel de poliacrilamida SDS-PAGE, al 10 %, BIS-TRIS (*Invitrogen*[®]), donde “<C” es la muestra antes de añadir la proteasa, “UL” se refiere a la proteasa *ULP* y “C” es la muestra resultantes de la digestión. Las muestras a la derecha del marcador de alto peso molecular (M) son la no retenida en la columna (NR), la lavada (L) y la eluida con el tampón de la Tabla III-C.



6.2 Caracterización de interacción BRMS1 aa51-84 - NMI aa16-64

Como se ha explicado anteriormente, tras obtener la estructura de BRMS1₅₁₋₈₄, nos centramos en intentar cristalizarlo en complejo con el motivo “coiled-coil” de NMI. Estudios preliminares obtenidos en el laboratorio, ya aludidos anteriormente en el apartado “I” de esta sección de *Resultados*, sugerían la implicación de la región amino-terminal de la proteína NMI en su interacción con la región también amino-terminal de BRMS1. Dicho dominio de NMI, comprende los aminoácidos 16 al 64. Con el fin de profundizar en el conocimiento de esta interacción e intentar aportar evidencias para la realización de futuros experimentos, se realizaron ensayos RMN y estudios calorimétricos o ITC.

6.2.1 Espectrometría de resonancia magnética nuclear (RMN)

Con un espectro bidimensional de RMN se puede hacer uso de las propiedades magnéticas de los protones (¹H) y de heteronúcleos (¹⁵N) con el fin de explorar el entorno químico que los rodea, que es distinto para cada uno de ellos en una misma molécula. Se obtiene un diagrama en dos dimensiones, donde en uno de los ejes se estudian los cambios

sobre los protones, y en el otro, sobre los heteronúcleos. Los movimientos en las señales obtenidas, son proporcionales a los cambios químicos que acontecen, de forma que si se produce, lo que se traduce en un desplazamiento de las señales en el espectro, siendo indicativo de la existencia de interacción.

Para abordar este experimento, necesitamos que una proteína contenga heteronúcleos, por lo que marcamos con ^{15}N el polipéptido BRMS1₅₁₋₈₄. Para tal fin, el vector de expresión pET28-SMT3-BRMS1₅₁₋₈₄ fue transformado en células de *E.coli* Rosetta(DE3), para su sobre-expresión con 0,5 mM IPTG a 37 °C, usando un medio mínimo complementado con NH_4Cl , como aporte de ^{15}N . Una vez sobre-expresada la proteína ^{15}N -His-SMT3-BRMS1₅₁₋₈₄, se purificó siguiendo el mismo protocolo que para proteína sin marcar (expuesto en el apartado “*Purificación de las construcciones con cola de histidinas*” en *Materiales y Métodos*). Ambos polipéptidos purificadas, NMI₁₆₋₆₄ y ^{15}N -BRMS1₅₁₋₈₄, fueron finalmente llevadas a una concentración de 300 μM y dializadas frente al tampón conteniendo 20 mM Fosfato Sódico pH 6,5; 50 mM NaCl y 1 mM DTT. Como control negativo de la interacción, utilizamos aprotinina, con un tamaño molecular de 6,5 kDa, similar al de la construcción NMI₁₆₋₆₄. Al igual que para las construcciones estudiadas de BRMS1 y NMI, la aprotinina fue dializada frente al mismo tampón y llevada a la misma concentración.

Los espectros fueron medidos usando un espectrómetro “Bruker AVANCE 700” a 25 °C de temperatura. Se recogieron tres tipos de espectros correspondientes a ^{15}N -BRMS1₅₁₋₈₄ individual, ^{15}N -BRMS1₅₁₋₈₄ - NMI₁₆₋₆₄ y ^{15}N -BRMS1₅₁₋₈₄ - Aprotinina como control negativo. Las señales no definidas del fragmento de ^{15}N -BRMS1₅₁₋₈₄ individual (Figura 50-A), sugiriendo una desestructuración de la construcción bajo estas condiciones. La inclusión del fragmento NMI₁₆₋₆₄ nos mostró un cambio importante en el patrón presentado por ^{15}N -BRMS1₅₁₋₈₄ (Figura 50-B), observándose señales bien definidas. Para visualizar dichas modificaciones, ambos espectros fueron superpuestos como se muestra en la Figura 50-C. En dicha superposición se observaron claramente desplazamientos químicos, con cambios en las posiciones y formas de las señales. este comportamiento, sugirió que existe interacción entre ambas construcciones, o expresado de otra forma; la presencia de NMI₁₆₋₆₄ hace cambiar la conformación de BRMS1₅₁₋₈₄.

Por otro lado, la superposición del espectro obtenido con ^{15}N -BRMS1₅₁₋₈₄ y el obtenido con ^{15}N -BRMS1₅₁₋₈₄ - Aprotinina (Figura 50-D), muestra que no existen cambios en el patrón de señales, por lo que ambas proteínas no interaccionan entre sí.

Con el fin de estudiar en grandes rasgos la especificidad de la unión, estos experimentos fueron repetidos disminuyendo 10 veces la concentración de ambos polipéptidos; es decir, hasta 30 μM . Tras analizar los datos obtenidos, observamos que los cambios químicos observados en el espectro al añadir el polipéptido NMI₁₆₋₆₄ a ^{15}N -BRMS1₅₁₋₈₄ son los mismos que los mostrados anteriormente. Por ello, se sugiere que la interacción entre ambos polipéptidos es específica.

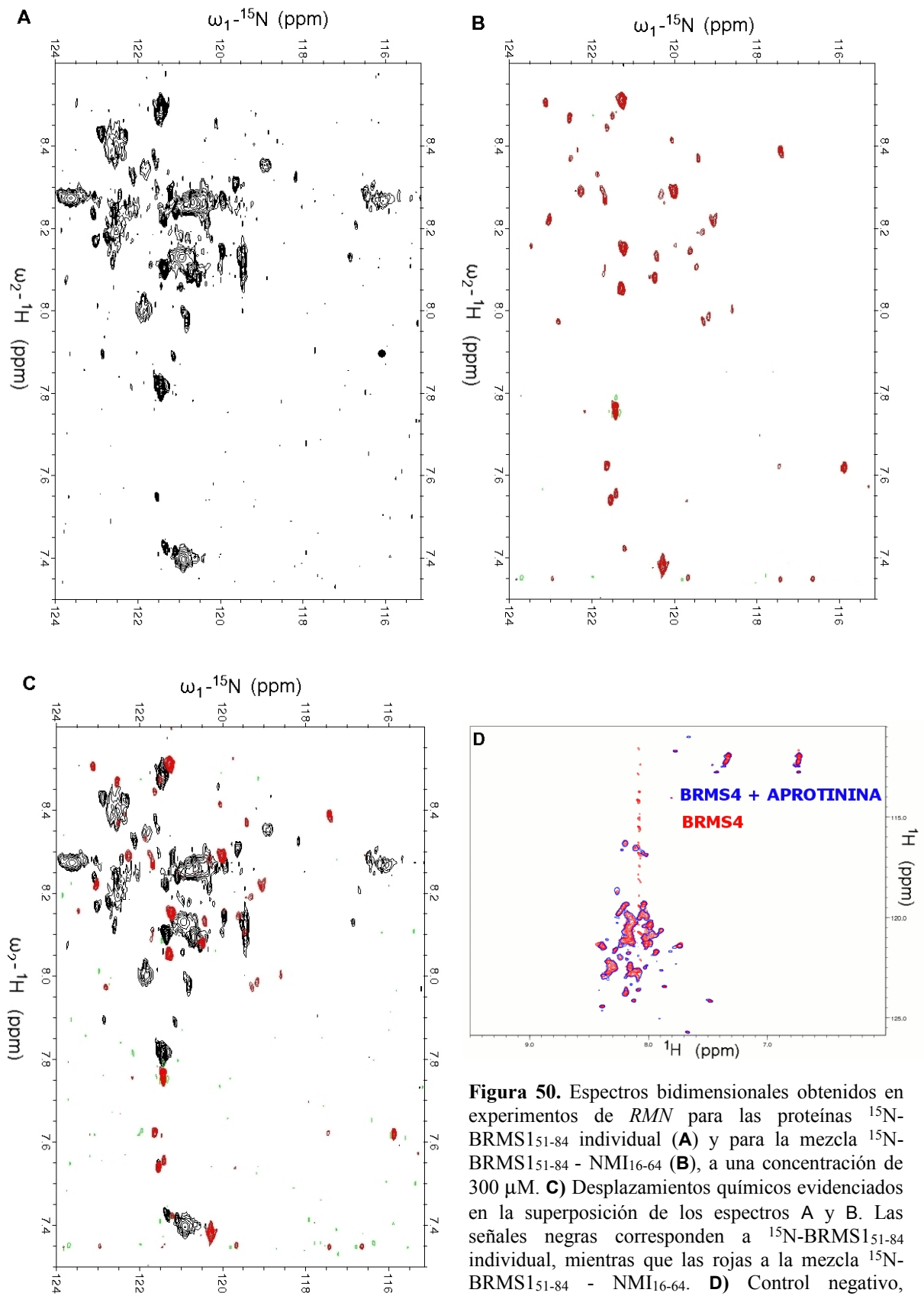


Figura 50. Espectros bidimensionales obtenidos en experimentos de RMN para las proteínas ^{15}N -BRMS151-84 individual (**A**) y para la mezcla ^{15}N -BRMS151-84 - NMI16-64 (**B**), a una concentración de 300 μM . **C**) Desplazamientos químicos evidenciados en la superposición de los espectros A y B. Las señales negras corresponden a ^{15}N -BRMS151-84 individual, mientras que las rojas a la mezcla ^{15}N -BRMS151-84 - NMI16-64. **D**) Control negativo, superposición de los espectros ^{15}N -BRMS151-84 y ^{15}N -BRMS151-84 - Aprotinina.

6.2.2 Calorimetría de titulación isotérmica (ITC)

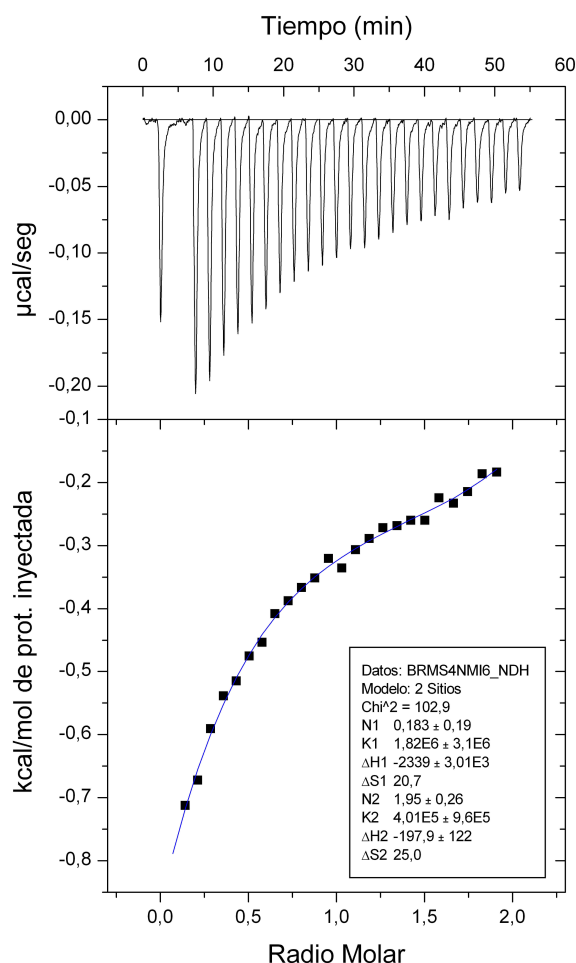
Tras comprobar la interacción entre BRMS1₅₁₋₈₄ - NMI₁₆₋₆₄ mediante ensayos de resonancia magnética, quisimos caracterizar esta unión en términos estequiométricos, y estimar los parámetros de afinidad. Para ello, realizamos ensayos de microcalorimetría donde se valora el calor liberado o absorbido en una reacción donde se añaden progresivamente pequeños volúmenes del ligando. En este caso se añadió el polipéptido BRMS1₅₁₋₈₄ sobre NMI₁₆₋₆₄ en solución.

Las proteínas purificadas fueron dializadas contra el mismo tampón conteniendo 10 mM HEPES pH 7,5 y 50 mM NaCl y posteriormente, las concentraciones ajustadas a 950 μM para BRMS1₅₁₋₈₄, situada en la jeringa y 95 μM para NMI₁₆₋₆₄, localizado en la celda. Tras las 25 inyecciones realizadas, a una temperatura de 20 °C, de BRMS1₅₁₋₈₄ sobre NMI₁₆₋₆₄, se produjo una diferencia de calor en el sistema y entre cada inyección, acercando al sistema a un estado de saturación. El mejor ajuste sobre la curva fue respondiendo a un modelo de interacción de dos sitios de unión (N1 y N2), con un valor de ajuste de Chi² de 102,9.

El efecto térmico neto de la reacción resultó lo suficientemente grande como para estimar valores de la entalpía, la constante de unión, y con ello la energía libre de Gibbs y la entropía. De este modo, los parámetros respecto a las ocupaciones de los sitios de unión resultaron ser de 0,183 para N1 y 1,95 para N2, con unas afinidades del orden de micromolar, siendo de 0,55 μM ($K1 = 1,82 \cdot 10^6 \pm 3,6 \cdot 10^6$ M) y de 2,5 μM ($K2 = 4,01 \cdot 10^5 \pm 9,6 \cdot 10^5$ M) respectivamente. El incremento de entropía para cada sitio de unión es de 20,7 J/K y de 25 J/K para N1 y N2 respectivamente. Debido a que dicho valor es positivo, y el término de incremento de entropía se refiere a la variación del orden molecular, los productos presentarían mayor desorden molecular que los reactivos. Al analizar los valores negativos de incremento de entalpía para ambos sitios de unión, $-2339 \pm 3 \cdot 10^3$ KJ/mol para N1 y -198 ± 122 KJ/mol para N2, observamos que se trata de una reacción exotérmica, donde se libera calor.

En la Figura 51 se muestra la diferencia de calor originada de cada inyección, así como los parámetros obtenidos del mejor modelo al que se ajustan los datos (de dos sitios de unión). Los datos procedentes de los dos ensayos: tampón inyectado sobre NMI₁₆₋₆₄ y BRMS1₅₁₋₈₄ inyectado sobre tampón, fueron sustraídos de la curva.

Figura 51. Calorimetría de Titulación Isotérmica para caracterizar la interacción de BRMS1₅₁₋₈₄ y NMI₁₆₋₆₄. La figura superior muestra los datos obtenidos durante la titulación, correspondiendo cada pico a la diferencia de calor de una inyección de la proteína de la jeringa (BRMS1₅₁₋₈₄, a 950 μ M) sobre la situada en la celda (NMI₁₆₋₆₄, a 95 μ M). El ajuste óptimo (representado con una línea) con un modelo de dos sitios de unión, se muestra en la figura inferior. Las afinidades de ambos sitios (N1 y N2) es de 0,55 μ M para N1 con una ocupación de 0,183 y 2,5 μ M para N2 con una ocupación de 1,95.



7. Ensayos de cristalización de los complejos caracterizados biofísicamente; BRMS1_{FL} + NMI_{FL} y BRMS1₅₁₋₈₄ + NMI₁₆₋₆₄

Tras llevar a cabo los experimentos expuestos anteriormente para comprobar las interacciones entre las construcciones BRMS1_{FL} y NMI_{FL} por un lado y BRMS1₅₁₋₈₄ + NMI₁₆₋₆₄ por otro, las muestras obtenidas mostrando la presencia de ambas proteínas, fueron concentradas hasta 9,2 mg/mL y 80 mg/mL respectivamente, para realizar ensayos de cristalización mediante difusión de vapor.

Con estas muestras en tampones conteniendo 20 mM TRIS pH 7,5 y 50 mM NaCl, para el complejo formado por las dos proteínas completas, y los respectivos tampones finales expuestos en los experimentos de RMN e ITC para los complejos formados por los motivos “coiled-coil”, se probaron alrededor de 1200 condiciones de cristalización, probando diferentes concentraciones de proteína y de temperatura de incubación. Finalmente, se obtuvieron cristales, de la muestra que presentaba las dos proteínas completas, como los que se exponen en la Figura 52, en diversas condiciones, también mostradas.

Los cristales surgieron tras 3 días, en gotas de 1,8 µL de volumen, con una proporción 1:1 entre proteína:reservorio y a 22 °C de temperatura.



Condiciones de cristalización	
	0,1 M CAPS pH 10,5; 30 % PEG 400
	0,1 M CHES pH 9,5; 30 % PEG 3000
	0,2 M MgCl ₂ ; 0,1 M Imidazol pH 8,0; 15 % EtOH
	0,1 M CHES pH 9,5; 30 % PEG 400
	0,2 M ZnAc ₂ ; 0,1 M Imidazol pH 8,0; 20 % 1,4-Butanediol
	0,2 M NaCl; 0,1 M CAPS pH 10,5; 20 % PEG 8000
	0,2 M MgCl ₂ ; 0,1 M TRIS pH 8,5; 3,4 M 1,6-Hexanediol
	0,2 M MgCl ₂ ; 0,1 M TRIS pH 8,5; 30 % PEG400
	0,1 M HEPES pH 7,5; 20 % Jeffamine M-600
0,2 M Mg Formate	0,1 M NaCitrate pH 5,6; 10 % Jeffam. M-600; 10 mM FeCl ₃
0,1 M TRIS pH 8,5; 0,2 M MgCl ₂ ; 30 % PEG 4000	0,1 M Bicina pH 9,0; 2 M MgCl ₂
0,1 M TRIS pH 8,5; 0,2 M LiSO ₄ ; 30 % PEG 4000	0,1 M TRIS pH 8,5; 8 % PEG 8000
0,1 M TRIS pH 8,5; 0,2 M NaAc; 30 % PEG 4000	0,1 M TRIS pH 8,5; 0,2 M NaCitrate; 30 % PEG 400

Figura 52. Cristales obtenidos de la muestra que presenta las proteínas completas BRMS1 y NMI, obtenida mediante co-expresión y co-purificación de las mismas. En la tabla se muestran las condiciones de cristalización donde surgieron cristales del tipo de la fotografía

Con el fin de descartar si estos cristales eran formaciones de sal o de proteína, se sometieron a difracción por rayos X en el sincrotrón de Grenoble, sin obtenerse ninguna reflexión en la recogida de los datos. El patrón obtenido no respondía a difracción típica de una red cristalina de moléculas salinas, lo que evidencia que se trata de cristales de proteína.

Se realizaron, sin éxito, numerosos intentos de optimización de los cristales a lo largo de todas las condiciones mostradas, con el fin de mejorarlos y poderlos someter a difracción por rayos X.

Capítulo 2

Tax1BP1.

Resumen

Tax1BP1, la proteína de unión a la oncoproteína Tax1 y envuelta en procesos de ubiquitinación, consta de dos dominios de unión a ubiquitinas (Iha *et al.*, 2008) cuya estructura se logró resolver en este proyecto. Además de mostrar la estructura de los dos dominios UBZs, se expondrán los resultados obtenidos en cuanto a clonaje, expresión y cristalización de las construcciones diseñadas, y diversos ensayos de caracterización de interacción con ubiquitina/s y otras proteínas como la ligasa de ubiquitina Itch.

Con ello, se abre un amplio abanico de posibilidades para continuar el estudio estructural y funcional de esta proteína en cascadas de ubiquitinación que regulan procesos celulares de gran relevancia.

1. Diseño de construcciones de la proteína Tax1BP1

La proteína *Tax1BP1* humana, se ha descrito que está involucrada principalmente en procesos de ubiquitinación, es una proteína de 789 aminoácidos que, además de contener tres motivos “coiled-coil”, una región homóloga a la proteína NDP52 y un dominio de oligomerización, presenta dos dominios “dedos de zinc” potenciales de unión a ubiquitina (Iha *et al.*, 2008), conocidos como *UBZs*, en su extremo carboxilo-terminal. Debido al interés que tienen ambos *UBZs* en procesos de ubiquitinación, y por lo tanto en la función global de *Tax1BP1* en el complejo multiproteico que forma con A20-Itch, que regula diversas moléculas de la ruta de activación del factor de transcripción NFκB a través de procesos de ubiquitinación, decidimos seleccionar estos dos dominios como objeto de estudios estructurales.

Con el fin de estudiar el grado de conservación evolutiva de la proteína completa en general y de estos dominios en particular, realizamos alineamientos entre los 9 ortólogos encontrados mediante *GeneCard*⁸ y *HomoloGene* (NCBI)¹². Los estudios (no mostrados) de porcentaje de homología y divergencia entre las secuencias proteicas (realizados con *MegAlign*, del programa *DNAStar*), evidencian que la mayor similitud es la que presentan la especie humana con el chimpancé (98,7 %) seguido del perro (88,5 %) y de la rata y el ratón (85,3 %).

El alineamiento de la secuencia aminoacídica de todas estas especies de la región de interés, que abarca los dos dominios potenciales de unión a ubiquitina (728-753 para *UBZ1* y 755-780 para *UBZ2*), mostró una elevada homología entre las distintas especies. Además, dos residuos de cisteínas de cada dominio (C730 y C733 en *UBZ1* y C757 y C760 en *UBZ2*) y dos de histidinas (H749 y H753 para *UBZ1* y H776 y H780 en *UBZ2*), involucrados en la unión al átomo de zinc del tipo Cys₂-His₂, se encuentran totalmente conservados, tal y como se muestra en la Figura 53, señalados con un asterisco.

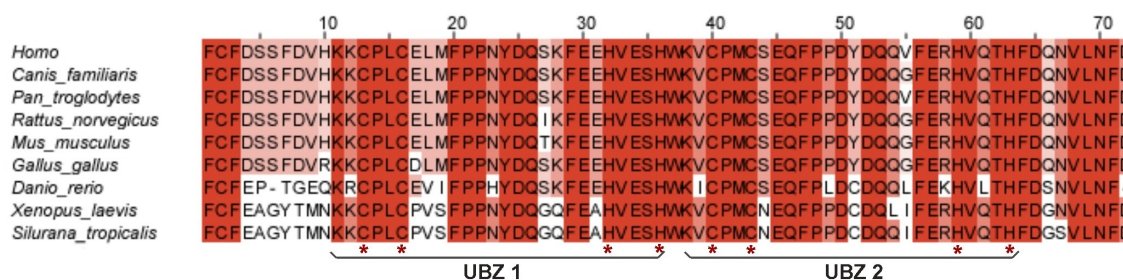


Figura 53. Alineamiento de la secuencia aminoacídica de los dos dominios de unión a ubiquitina (*UBZ1* y *UBZ2*) de la proteína *Tax1BP1* entre diferentes especies (*Jalview*).

Se muestran en rojo los aminoácidos totalmente conservados y progresivamente, en colores mas suaves, los residuos que están menos conservados. En blanco se exponen los aminoácidos que no están conservados. Los residuos que están envueltos en la unión a átomos de zinc (histidinas y cisteínas) están marcados con un asterisco.

Con ayuda de la predicción de estructura secundaria proteica predicha (*PSIPRED*¹¹, Figura 54), intentamos evitar regiones no estructuradas y respetar las que mostraban estructura definida. Con estas premisas, realizamos el diseño de tres construcciones de

Tax1BP1 en torno a los dos dominios potenciales de unión a ubiquitina: *UBZ1* (aa 725-756), *UBZ2* (aa 755-789) y *UBZ1+2* (aa 725-789).

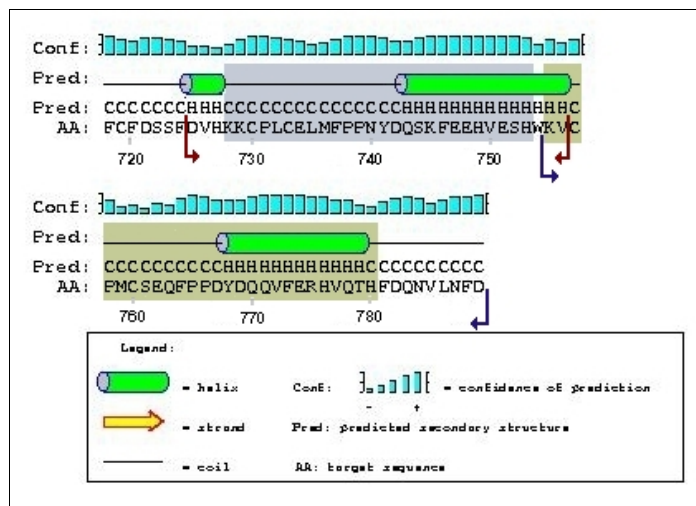


Figura 54. Predicción de estructura secundaria proteica de los dos dominios *UBZs* de *Tax1BP1*. El dominio *UBZ1* predicho abarca la región encuadrada en azul (del residuo 728 al 753), y la caja marrón el *UBZ2* (del 755 al 780). El segmento clonado para el *UBZ1* fue el comprendido entre las flechas rojas, mientras que el *UBZ2*, el flanqueado por las flechas azules. La predicción representada con un cilindro verde corresponde a estructura en α -hélice (H), mientras que la flecha amarilla y línea negra (C) a β -lámina y región sin estructura definida, respectivamente. En la primera fila se muestra la confiabilidad en la predicción, que es mayor a barras más elevadas.

2. Expresión, purificación, cristalización y resolución de la estructura de dominios de unión a Ubiquitina (*UBZs*) de *Tax1BP1*

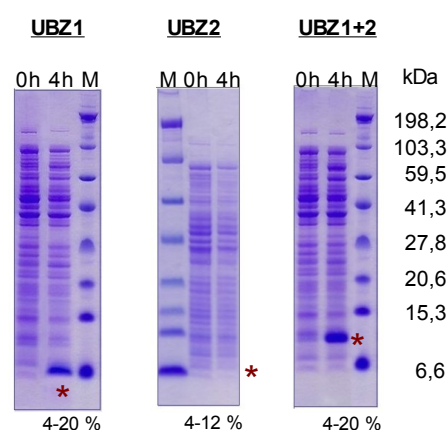
Las tres regiones de interés que engloban los dominios predichos de unión a ubiquitina diseñadas; *UBZ1* (aa 725-756), *UBZ2* (aa 755-789) y *UBZ1+2* (aa 725-789); se amplificaron y clonaron en un vector de expresión derivado de pET28 (PP-pET28). Dicha modificación añade seis histidinas en el extremo amino-terminal, que son eliminadas tras la digestión con la proteasa *PreScission*, originando una secuencia que añade los residuos GPHM en dicho extremo.

Tras transformar estos plásmidos en células de *E.coli*, y realizar numerosos ensayos de optimización de su expresión, se obtuvieron elevados niveles de sobre-expresión del dominio 1 de unión a ubiquitina (*UBZ1*) y de los dos dominios unidos (*UBZ1+2*), usando la cepa *Rosetta(DE3)* y el medio de crecimiento *2xYT*, e induciendo la expresión con 1 mM de IPTG, a 37 °C durante 4 horas. A pesar de que se realizaron numerosos intentos de expresión sobre el dominio *UBZ2* individual, además de probar otros vectores, medios de crecimiento y cepas de expresión, no se consiguieron niveles de proteína de fusión sobre-expresada adecuados para abordar ensayos de cristalización. En la Figura 55, se muestran los niveles de sobre-expresión obtenidos para las tres construcciones.

Figura 55. Expresión de His-PP-*UBZ1*, His-PP-*UBZ2* y His-PP-*UBZ1+2* en cepa de *E.coli*, donde “PP” alude al sitio de reconocimiento de la proteasa *PreScission*. Las muestras se analizaron mediante electroforesis en gel de poliacrilamida (SDS-PAGE) de diferentes porcentajes de acrilamida BIS-TRIS, referidos bajo los geles.

El carril marcado como “0h” se corresponde con el total de proteínas expresadas antes de inducir la sobre-expresión de la proteína heteróloga mediante adición de IPTG. El carril “4h” contiene el total de proteínas presente en las células transcurridas 4 horas de añadir el agente inductor.

Para referenciar los tamaños moleculares se usó un marcador (M) de alto peso molecular (*BioRad*^E), cuyos pesos moleculares se muestran a la derecha en unidades de kDa.



El asterisco rojo indica la banda correspondiente a la proteína de interés sobre-expresada. Los tamaños moleculares son: *UBZ1*: 6,4 kDa; *UBZ2*: 6,8 kDa y *UBZ1+2*: 10,5 kDa.

Las dos proteínas que se lograron sobre-expresar con éxito (*UBZ1* y *UBZ1+2*), fueron purificadas y sometidas a numerosos ensayos de cristalización. Aunque ambas construcciones se consiguieron cristalizar (cristales de escasa calidad referentes a *UBZ1*, no mostrados), finalmente logramos resolver la estructura del *UBZ1+2*, y debido a que este fragmento engloba al *UBZ1*, se expondrán únicamente los resultados, en cuanto a purificación y cristalización, obtenidos con *UBZ1+2*, que son extrapolables al *UBZ1*.

2.1 Purificación de *UBZ1+2* (aa 725-789)

El cultivo celular procariota resultante de la expresión de His-PP-*UBZ1+2* (“PP” representa el sitio de reconocimiento de la proteasa *PreScission*), fue lisado mediante sonicación, con un tampón con 50 mM TRIS pH 7,5 y 0,5 M NaCl, complementando con 0,5 % Tween-20 (v/v) con el fin de favorecer la solubilización de la proteína sobre-expresada. Posteriormente dicha muestra fue centrifugada con la idea de extraer la masa proteica y someterla a diferentes etapas de purificación. Tras la centrifugación se obtuvieron dos fases bien diferenciadas, una soluble, donde se localizó la proteína de interés, y otra insoluble (“S” e “I” respectivamente), que se exponen en la Figura 56.

2.1.1 Cromatografía de afinidad de His-PP-*UBZ1+2*

El sobrenadante obtenido tras la lisis y ultracentrifugación (“S” en la Figura 56) se filtró e inyectó en una columna de afinidad a histidinas, precargada con níquel. Tras ser lavada con el tampón de la Tabla III-B, conteniendo 10 mM Imidazol, se aplicó un gradiente lineal del mismo hasta una concentración de 0,5 M con el fin de eluir las distintas proteínas retenidas en la columna según su afinidad al níquel, que se colectaron en fracciones de 4 mL.

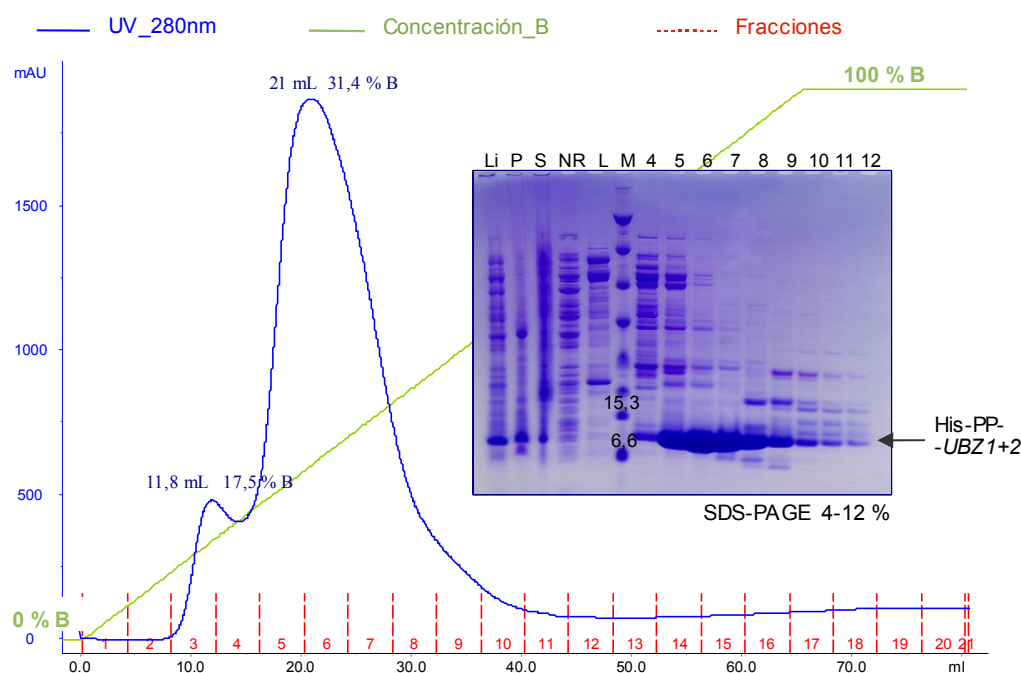


Figura 56. Cromatograma y muestras resultantes de la columna de afinidad a histidinas en una purificación de His-PP-UBZ1+2. El cromatograma muestra la absorbancia a una longitud de onda de 280 nm, de la mezcla proteica soluble cargada en la columna, cuyos componentes son eluidos al aplicar un gradiente continuo de 0 % a 100 % del tampón de elución (B), correspondiendo desde 10 mM hasta 0,5 M de Imidazol. Estas muestras se analizaron mediante electroforesis en gel de poliacrilamida (SDS-PAGE) de 4-12 %, BIS-TRIS. Los carriles expuestos en el mismo hacen referencia a: “Li”: Lisado celular; “I” y “S”: fracciones insoluble y soluble respectivamente, tras la lisis celular y centrifugación; “NR”: muestra no retenida en la columna; “L”: lavado; “M”: marcador de alto peso molecular, donde se representa el tamaño en kDa de los dos pesos limitantes de la banda de interés; “4-12”: fracciones eluidas (mostradas en rojo en el cromatograma) tras aplicar el gradiente de Imidazol. El tamaño de la proteína de fusión His-PP-UBZ1+2 es de aproximadamente 10,5 kDa.

Las fracciones 5 a 8 que presentaban la proteína de interés, eluida a una concentración de imidazol de 157 mM, fueron unidas para someterlas a digestión enzimática. Es preciso mencionar que aunque se intentó optimizar la purificación a través de esta primera columna, mediante sucesivas etapas de lavado incrementando la concentración de imidazol, así como utilizando diferentes protocolos, no conseguimos disminuir la presencia de proteínas contaminantes.

2.1.2 Proteolisis y cromatografía de exclusión por tamaño molecular

Con el fin de eliminar la cola de histidinas en la construcción His-UBZ1+2, añadimos a la muestra 80 µg de proteasa *PreScission* (1 µg de proteasa por cada mg de proteína a digerir). La proteolisis se llevó a cabo dializando frente a un tampón de 50 mM TRIS pH 7,5 y 0,2 M NaCl, a temperatura ambiente y durante 90 minutos, que resultaron suficientes para digerir la proteína de fusión prácticamente en su totalidad. Los productos de digestión originados se muestran en la Figura 57-A (carril nombrado como “C”). Esta muestra fue concentrada hasta un volumen de 1,2 mL e inyectada en una columna de exclusión por tamaño molecular (“Superdex 75 16/60”). El patrón de absorbancia

obtenido a una longitud de onda de 280 nm (Figura 57-B), mostró un máximo mayoritario, correspondiente a las fracciones 20-22. Dichas fracciones fueron analizadas mediante electroforesis en gel de poliacrilamida, teñido con azul de Coomassie, mostrando la presencia de la proteína de interés.

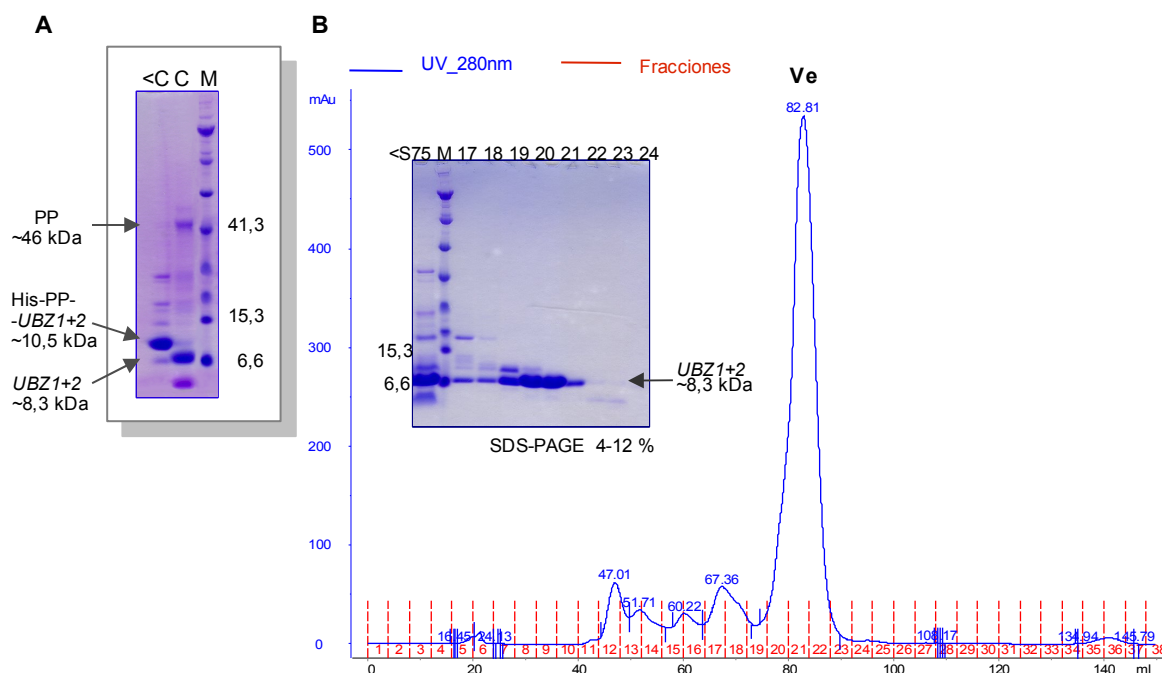


Figura 57. Muestras resultantes de la proteólisis de His-PP-UBZ1+2 (A) y cromatografía de exclusión por tamaño molecular de la proteína digerida (B), analizadas mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 4-12 % BIS-TRIS (*Invitrogen*^l).

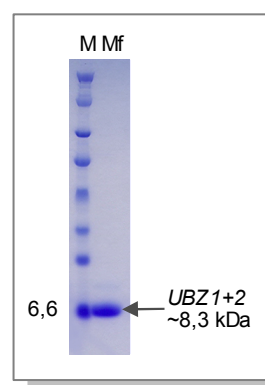
Los carriles “<C” y “C” se corresponden con muestras tomadas de la mezcla proteica antes y después de la proteólisis respectivamente. La banda referente a la proteasa PeScission se representa como “PP”.

En el cromatograma se muestra la absorbancia a una longitud de onda de 280 nm, de las proteínas eluidas en distintos volúmenes según su tamaño y recogidas en fracciones de 4 mL representadas en rojo. Se muestra sobre cada pico de elución, su correspondiente volumen, siendo “Ve” el de la proteína de interés. Las muestras antes de la inyección en la columna “Superdex 75 16/60” (<S75) y eluidas en el pico mayoritario (19-22) se muestran en el gel de poliacrilamida. “M”: marcador de alto peso molecular, donde se representa el tamaño en kDa de los pesos cercanos a las bandas de interés en ambos geles.

El peso molecular del volumen de elución correspondiente al máximo de elución de la proteína (Ve = 82,81 mL), fue estimado usando una curva de calibración obtenida con estándares conocidos. El tamaño obtenido aproximado fue de 10,7 kDa; lo que sugiere un comportamiento monomérico (peso molecular teórico de UBZ1+2 es 8,3 kDa).

Las fracciones 20, 21 y 22 fueron concentradas hasta 60 mg/mL. Tras el proceso de purificación, se obtuvo una proteína final que presentó un elevado grado de pureza, como se puede visualizar en el gel de la Figura 58, teñido con una solución de azul de “Coomassie”.

Figura 58. Gel de poliacrilamida del 4-12 % BIS-TRIS, mostrando un elevado grado de pureza de la proteína final *UBZ1+2* (muestra final; “Mf”). Se usó un marcador (“M”) de estándares de alto peso molecular. Se muestra el tamaño molecular en kDa de la banda próxima a la proteína de interés.



Tras la digestión con la proteasa específica *PreScission*, se originó una proteína de 69 aminoácidos, de los cuales, únicamente 4 residuos son añadidos de forma exógena en su región amino-terminal, según se muestra en la Figura 59. Este es el fragmento final usado en los ensayos de cristalización.

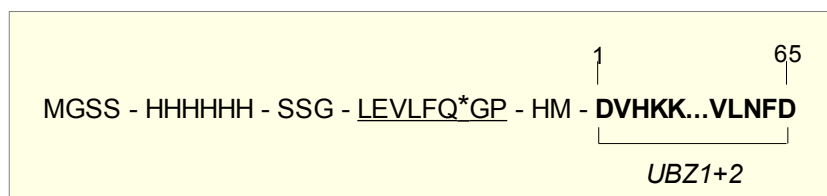


Figura 59. Secuencia aminoacídica de la construcción His-PP-*UBZ1+2*, comprendiendo los aminoácidos 725 al 789 de la proteína Tax1BP1, tras su expresión. Con un asterisco se muestra el sitio de corte de la proteasa *PreScission* cuya secuencia de reconocimiento se presenta subrayada. Los residuos que se muestran en “negrita” son los correspondientes a los dominios *UBZ1+2*, concretamente a los comprendidos entre los aminoácidos 725 (aspártico 1 en la figura) y 789 (aspártico 65 en la figura) de Tax1BP1. El fragmento final generado tras la proteólisis, añade cuatro residuos extras (GPHM) en su extremo amino-terminal.

Tras su proteólisis, la proteína *UBZ1+2*, se concentró hasta 60 mg/mL en un tampón 20 mM TRIS pH 7,5 y 0,2 M NaCl, y se utilizó para realizar ensayos de cristalización. Para ello se siguió el método de difusión de vapor, mediante gota sentada en placas de 96 pocillos, usando inicialmente diversos bloques comerciales. Tras un proceso de optimización, que consistió en probar alrededor de 800 condiciones de cristalización, diferentes volúmenes de gota, así como de concentraciones y variar la temperatura de incubación, se obtuvieron cristales, transcurrido un día desde el inicio del experimento. Estos cristales iniciales, con forma cúbica y de unas dimensiones aproximadas de 0,03 x 0,03 x 0,03 mm (Figura 60-A), aparecieron en una condición con 0,1 M HEPES pH 7,5 y 1,5 M Sulfato de Litio, a una temperatura de 22 °C y en un volumen de gota de aproximadamente 1,6 µL (con una proporción equivolúmetrica entre proteína y reservorio).

En el proceso de optimización de los cristales que acabamos de describir, obtuvimos dos tipos de morfologías diferentes, variando únicamente el pH, de forma que al disminuirlo hacia 7,0-7,2 se favorecía una morfología elongada, y al aumentarlo hacia 8,0-8,5 se favorecía una forma cúbica (Figura 60-B). Los primeros cristales no difractaron bajo un haz de Rayos X del sincrotrón, sin embargo, las estructuras cúbicas desembocaron en cristales de mejor calidad, en una condición de cristalización de 0,1 M TRIS pH 8,5 y 1

M Sulfato de Litio. Estos cristales optimizados, mostrados en la Figura 60-C, y de un tamaño aproximado de 0,8 x 0,7 x 0,3 mm, pudieron ser usados para la obtención de un juego completo de datos de difracción.

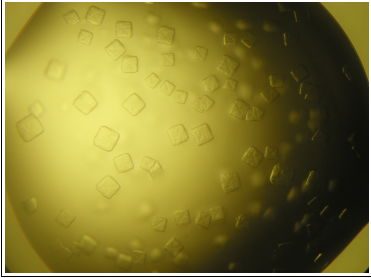
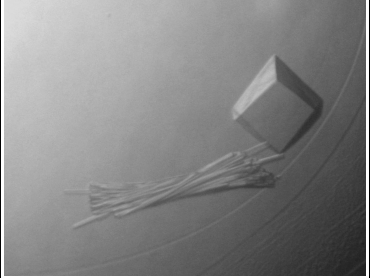
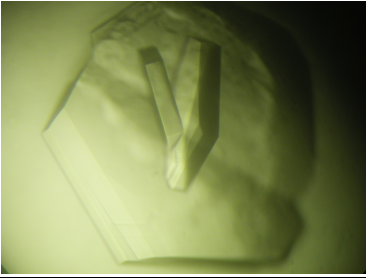
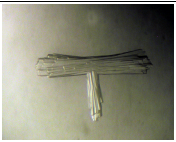
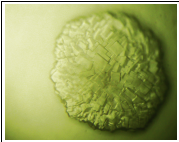
A	B		C
			
			
Condición comercial: 0,1 M HEPES pH 7,5 1,5 M Sulfato de Litio	pH 7,2 (HEPES)	pH 8,5 (TRIS)	0,1 M TRIS pH 8,5 1 M Sulfato de Litio

Figura 60. Cristales de *UBZ1+2* obtenidos mediante difusión de vapor, a una concentración de 60 mg/mL, usando una proporción 1:1 (proteína:reservorio) y a 22 °C de temperatura. **A)** Cristales iniciales, obtenidos bajo la condición comercial expuesta, en gota sentada y formato de 96 pocillos. **B)** Cristales optimizados, en gota colgante, y formato de 24 pocillos. A pH más ácido se favorece la formación de cristales con morfología elongada, mientras que a pH más básico, surgen formas cúbicas. **C)** Cristales optimizados bajo la condición mostrada, en gota colgante, usado para experimentos de difracción y resolución de la estructura.

2.2 Estructura de *UBZ1+2*

El cristal mostrado en la Figura 60-C fue impregnado en Paratone-N con el fin de eliminar el exceso de solvente que pudiese producir difracción en forma de aros de hielo, y finalmente se montó en un lazo de 0,4-0,5 mm, para ser rápidamente enfriado a 130 K y difractado con un haz de rayos X, en la línea ID14-1 del sincrotrón (Grenoble), usando una longitud de onda de 0,934 Å.

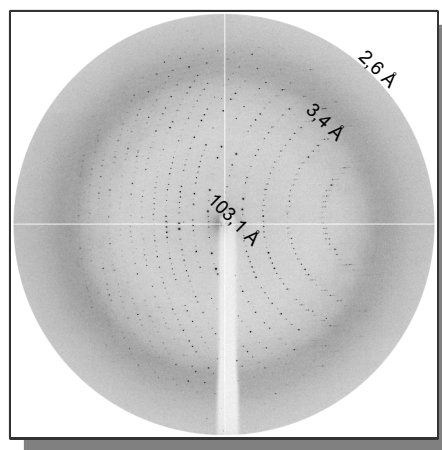
Para coleccionar un juego de datos completamente informativo, se estudió, con ayuda de la herramienta *Strategy* de *Mosflm*, la estrategia a seguir, resumida en la Tabla XVII. Con ello, establecimos una distancia entre el cristal y el detector de aproximadamente 277 mm, alcanzando una resolución máxima de 2,59 Å. Se osciló 0,25 ° por cada imagen hasta coleccionar un total de 110 °.

Energía	Distancia	Resolución	Incremento del ángulo ϕ total	Incremento de ϕ / Imagen	Tiempo de exposición	Número de pases
0,934 Å (13,274 keV)	277,26 mm	103,14 - 2,59 Å	110 °	0,25 °	6 seg.	3

Tabla XVII. Parámetros usados en la recolección del conjunto de datos de difracción de un cristal de *UBZ1+2* en la línea ID14-1 del sincrotrón. Se muestran por orden, la longitud de onda y energía con la que se emite el haz de rayos X incidente sobre el cristal; la distancia entre el cristal y el detector y los límites de resolución entre los que se recogieron las difracciones emitidas por las moléculas del cristal. A continuación se muestran parámetros del ángulo “phi” (ϕ) que determina el posicionamiento del goniómetro donde se soporta el cristal. Estos son el incremento del ángulo total colectado y el incremento en la rotación del mismo en cada imagen. También se expone el tiempo de exposición en segundos para la toma de cada imagen y el número de repeticiones para cada imagen.

Finalizada la recolección, se obtuvo un patrón de difracción como el que se expone en la Figura 61.

Figura 61. Imagen representativa del patrón de difracción obtenido de un cristal de *UBZ1+2*, obtenido en la línea ID14-1 del sincrotrón. Cada imagen fue rotada 0,25 ° durante 6 seg., recogiendo finalmente un total de 110 °. Sobre estas figuras se muestran algunos límites de resolución, expresados en angstroms (Å).



Tras la reducción y escalado de estos datos, usando para ello los programas *Mosflm* y *Scala* de *CCP4*, obtuvimos unas longitudes de los ejes de la celda de: $a = 118,88$ Å; $b = 118,88$ Å y $c = 327,35$ Å; con unos ángulos $\alpha = 90^\circ$; $\beta = 90^\circ$, $\gamma = 120^\circ$. El grupo espacial al que respondía la simetría del cristal resultó ser *P6₅22*, determinado con ayuda del programa *Pointless* (del paquete *CCP4*). Los parámetros obtenidos tras este procesamiento de los datos se resumen en la Tabla XVIII.

Tax1BP1 aa (725-789)	
Reducción de datos	
Grupo espacial	<i>P</i> 6 ₅ 22
Dimensiones de la celda	
a; b; c (Å)	118,88; 118,88; 327,35
α , β , γ (°)	90; 90; 120
Rango de resolución (Å) (última capa)	51,50 - 2,59 (2,73 - 2,59)
Observaciones (únicas)	527791 (43605)
Completitud (%) (última capa)	100 (99,9)
Multiplicidad (última capa)	12,1 (10,4)
R _{sym} ^a (última capa)	0,095 (0,793)
R _{meas} ^b (última capa)	0,108 (0,870)
R _{pim} ^c (última capa)	0,030 (0,264)
$\langle I/\sigma \rangle$ (última capa)	19,1 (3,1)
Completitud anómala (%)	100
Multiplicidad anómala	6,5
DelAnom correlation between half-sets	0,363
Mid-Slope of Anom Normal Probability	1,213

^a R_{sym} es el *factor R* no sopesado sobre las intensidades simétricamente relacionadas.

$R_{sym} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, donde $I_i(hkl)$ es la intensidad de la reflexión hkl y $\langle I(hkl) \rangle$ es la intensidad media de todas las observaciones i de la reflexión hkl .

^b $R_{meas} = \{\sum_{hkl} [N/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|\} / \sum_{hkl} \sum_i I_i(hkl)$, donde $I_i(hkl)$ son las intensidades observadas, $\langle I(hkl) \rangle$ son las intensidades medias y N es la multiplicidad de la reflexión hkl .

^c $R_{pim} = \{\sum_{hkl} [1/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|\} / \sum_{hkl} \sum_i I_i(hkl)$, donde $I_i(hkl)$ son las intensidades observadas, $\langle I(hkl) \rangle$ son las intensidades medias y N es la multiplicidad de la reflexión hkl .

$\langle I/\sigma \rangle$ es la media de la relación entre la intensidad de la difracción y el ruido de fondo.

Tabla XVIII. Datos estadísticos cristalográficos obtenidos tras el procesamiento y simplificación de los datos de difracción recogidos del cristal de UBZ1+2 de la proteína Tax1BP1.

Los análisis realizados *a priori*, con la herramienta *Matthews_Coef* de *CCP4*, para determinar la cantidad de solvente en el cristal y estimar el número de moléculas en la unidad asimétrica, sugirieron diferentes coeficientes y diversas posibilidades de número de moléculas en la unidad asimétrica (Tabla XIX). De todas las opciones ofrecidas, aquellas con la probabilidad mas elevada (superior a 0,08), y con un coeficiente de entre 2 y 3 Å³/Da son las que están en el rango de entre 14 y 20 moléculas por unidad asimétrica, que corresponderían a un porcentaje de solvente de entre 57,11 y 38,72 % respectivamente (Kantardjieff y Rupp, 2003; Matthews, 1968).

Nº molec/ UA	Coef. de Matthews (Å ³ /Da)	Solvente (%)	P(2,59)	P(Tot)	Nº molec/ UA	Coef. de Matthews (Å ³ /Da)	Solvente (%)	P(2,59)	P(Tot)
13	3,09	60,17	0,05	0,05	18	2,23	44,85	0,16	0,15
14	2,87	57,11	0,08	0,08	19	2,11	41,79	0,12	0,12
15	2,67	54,04	0,11	0,11	20	2,01	38,72	0,07	0,08
16	2,51	50,98	0,14	0,14	21	1,91	35,66	0,03	0,04
17	2,36	47,91	0,16	0,16					

Tabla XIX. Coeficientes de *Matthews* y probabilidades normalizadas para cada posible número de moléculas en la unidad asimétrica (UA) y porcentaje de solvente para el cristal procesado de *UBZI*+2. Estimación facilitada por *Matthews-Coef* (de *CCP4*), que sugiere entre 14 y 20 moléculas por unidad asimétrica. P(2,59) alude a la probabilidad usando el límite de resolución de 2,59 Å, y P(Tot) a la probabilidad utilizando todos los rangos de resolución.

Debido a la ausencia de un modelo para poder utilizarlo en la resolución de la estructura mediante Reemplazo Molecular, intentamos obtener un cristal derivado de átomo pesado. Tras múltiple intentos, no obtuvimos un cristal derivado de mercurio, oro o platino que nos llevase a la obtención de un juego de datos procesable. Sin embargo, pudimos utilizar la señal anómala procedente de los átomos de zinc, coordinados en los dominios “dedos de zinc” *UBZI* y *UBZ2*, de un cristal sin tratar. La longitud de onda usada en la recolección fue de 0,934 Å.

Con el fin de encontrar las fases y con ello intentar construir un modelo inicial, se utilizó dicho conjunto de reflexiones integradas y escaladas con señal anómala y la secuencia de *UBZI*+2. Sin imponer número de cadenas por unidad asimétrica, y utilizando *autoSHARP*, se obtuvieron, por *SAD*, las fases experimentales y se encontraron 22 posiciones de átomos pesados (Zn), todos ellos con una ocupación de mas del 78 %. Dichas coordenadas fueron usadas para, junto con las fases obtenidas, construir un modelo usando *Resolve* de *PHENIX* (Autobuild). Para facilitar el entendimiento del procedimiento llevado a cabo, se ha confeccionado un esquema (Figura 62) que engloba la estrategia utilizada para resolver la estructura.

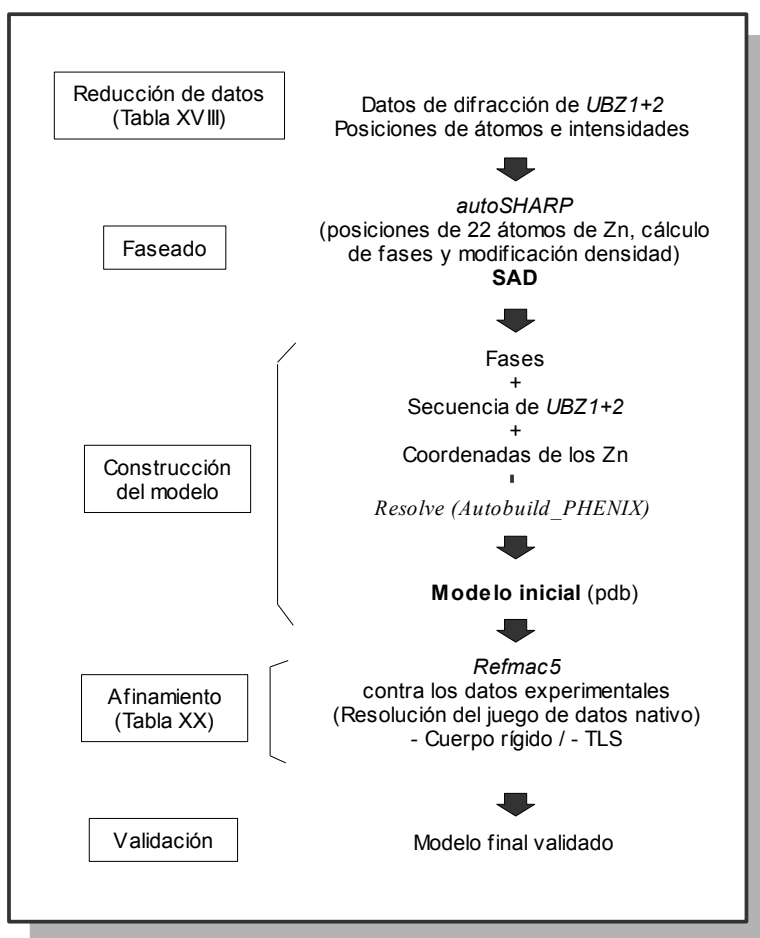


Figura 62. Estrategia utilizada para resolver la estructura de *UBZ1+2* de Tax1BP1 mediante *SAD*.

En cursiva se muestran los programas usados en cada caso.

Tras el modelado realizado por *PHENIX*, se obtuvo un modelo inicial que incluía 11 moléculas en la unidad asimétrica (*A-K*), que se repiten a lo largo del cristal formando un entramado organizado tal y como se puede observar en la Figura 63. Se puede tener una visión general de dicha disposición observando dos orientaciones diferentes de la celda, capturadas desde dos ángulos distintos. Una de ellas, se muestra a través del eje *a* (panel 63-A), donde se observan conformaciones organizadas en forma de rosetas contactadas entre sí para formar el cristal. La otra orientación, a través del eje *c* (panel 63-B), muestra mas claramente los canales de solvente que atraviesan el cristal, correspondiendo a un porcentaje 66,3 %.

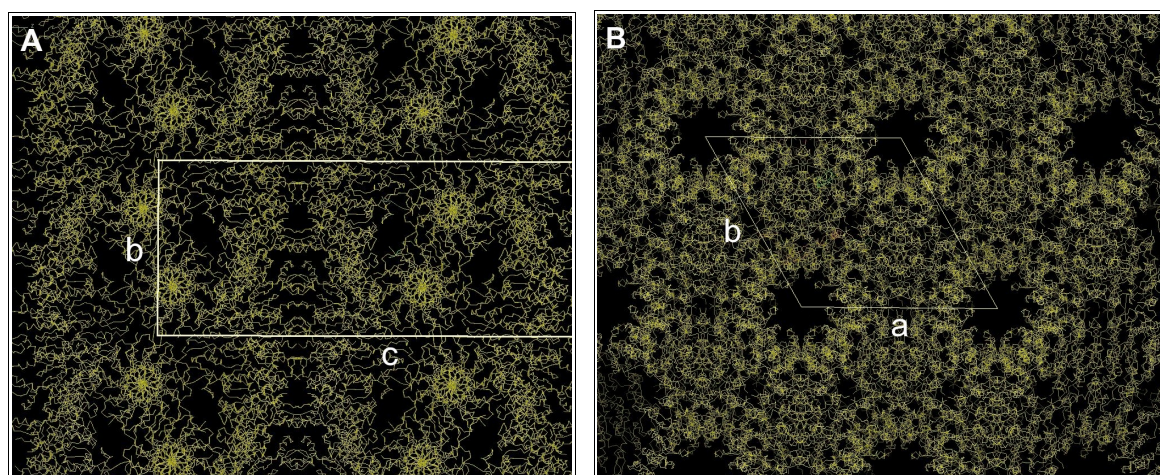


Figura 63. Empaquetamiento en el cristal de las once cadenas de los dos dominios *UBZ* de Tax1BP1 que conforman la unidad asimétrica. Se muestran dos orientaciones diferentes de la celda; **A)** a través del eje *a*; es decir, con dicho eje dispuesto de forma perpendicular a dicha vista, y **B)** a través del eje *c*. Cabe mencionar que la orientación a través del eje *a* genera una visión similar a la del eje *b*.

Tras un proceso de afinado con *Refmac5* (*CCP4*), en el que se limitó la resolución máxima en 2,75, se utilizaron 11 grupos de TLS (uno por cadena), y se aplicó simetría no cristalográfica (NCS) de los residuos 16 a 60 de la cadena *C* sobre las diez restantes (las restricciones usadas fueron de carácter medio tanto para las cadenas principales como para las laterales), se obtuvo un modelo final, cuyos parámetros se reflejan en la Tabla XX.

Tax1BP1 aa (725-789)	
Afinamiento	
Nº átomos no hidrógenos (moléculas de agua)	5887 (92)
Heteroátomos	
Cl	12
Zn	22
R^b (R_{free}^c)	0,215 (0,255)
$R.m.s.d.^d$ enlaces (Å)	0,010
$R.m.s.d.^d$ ángulos (°)	1,346
$Factor B^e$ (Å ²)	
Proteína	26,90
Solvente	14,71
Heteroátomos	20,25

^b $R = \sum_{hkl} \{ [F_{obs}(hkl)] - [F_{calc}(hkl)] \} / \sum_{hkl} [F_{obs}(hkl)]$, donde F_{obs} y F_{calc} son los factores de estructura observados y calculados respectivamente.

^c R_{free} es un parámetro de validación, correspondiente al *factor R* del 5 % de reflexiones que no han sido introducidas en el proceso de afinamiento.

^d $R.m.s.d.$ del inglés “Root mean square deviation”, es la desviación cuadrática media de las longitudes.

^e *Factor B* es un parámetro de estimación de la vibración del átomo con respecto a su posición central

Tabla XX. Datos estadísticos cristalográficos obtenidos tras el proceso de afinamiento del modelo final obtenido de *UBZ1+2*, realizado con *Refmac5* (*CCP4*).

Como se muestra en la Tabla XX, los factores estimados que determinan el ajuste del modelo; es decir, el *factor R* total y el correspondiente al 5 % de reflexiones que no se utilizaron en el proceso de afinamiento (*R_{free}*), fueron finalmente de 21,5 % y de 25,5 % respectivamente. Además del modelo ocupado por las once cadenas de aminoácidos, se modelaron 22 de zinc, 92 moléculas de agua y 12 átomos de cloro.

De los 759 residuos totales (69 aminoácidos por cadena) que componen la unidad asimétrica, modelamos 677, representando la gran mayoría de ellos. En la Tabla XXI, se muestra el número de residuos construidos de cada cadena. Las moléculas mas completas en sus extremos carboxilo- y amino- terminales fueron respectivamente la *C* y la *H*, que presentan cada una los residuos 2-67 y 1-64 respectivamente.

A	B	C	D	E	F	G	H	I	J	K
4-65	6-67	2-67	7-64	6-65	8-65	6-64	1-64	2-64	2-65	6-66

Tabla XXI. Residuos modelados para las 11 cadenas (A-K) que componen la unidad asimétrica del cristal de *UBZI*+2. En azul y en verde se muestran respectivamente las cadenas mas elongadas en sus extremos carboxilo- y amino- terminales.

En la Figura 64 se puede observar la disposición de estas 11 cadenas en la unidad asimétrica. Cada cadena se localiza entre dos colindantes, en planos completamente paralelos, y dispuestas de forma antiparalela con respecto a ellas según su eje longitudinal. Por este motivo se localizaron dos grupos de cadenas manteniendo el mismo sentido, es decir, con los extremos amino- y carboxilo- terminales apuntando hacia el mismo lado. Estos dos grupos lo forman *A-C-E-G-I-K* por un lado y *B-D-F-H-J* por otro, de forma que la cadena *A* está girada sobre *C* unos 32,6 °; 64,7 ° con respecto a *E*; 97,7 ° con *G*; 131 ° con *I* y 163,7 ° con *K*. Del mismo modo, los ángulos de rotación observados de la cadena *B* con respecto a *D* es 31,5 °; 64,5 con *F*; 97,3 ° con *H* y 131,2 ° con respecto a *J*. Las imágenes mostradas en la Figura 64, se corresponden con dos orientaciones diferentes de la celda, de forma que la primera es una vista de las once cadenas a través del eje *C*, donde se observa la disposición paralela entre ellas y la segunda a través del eje *A*, donde se muestra el giro en sentido horario de las cadenas dentro de cada grupo.

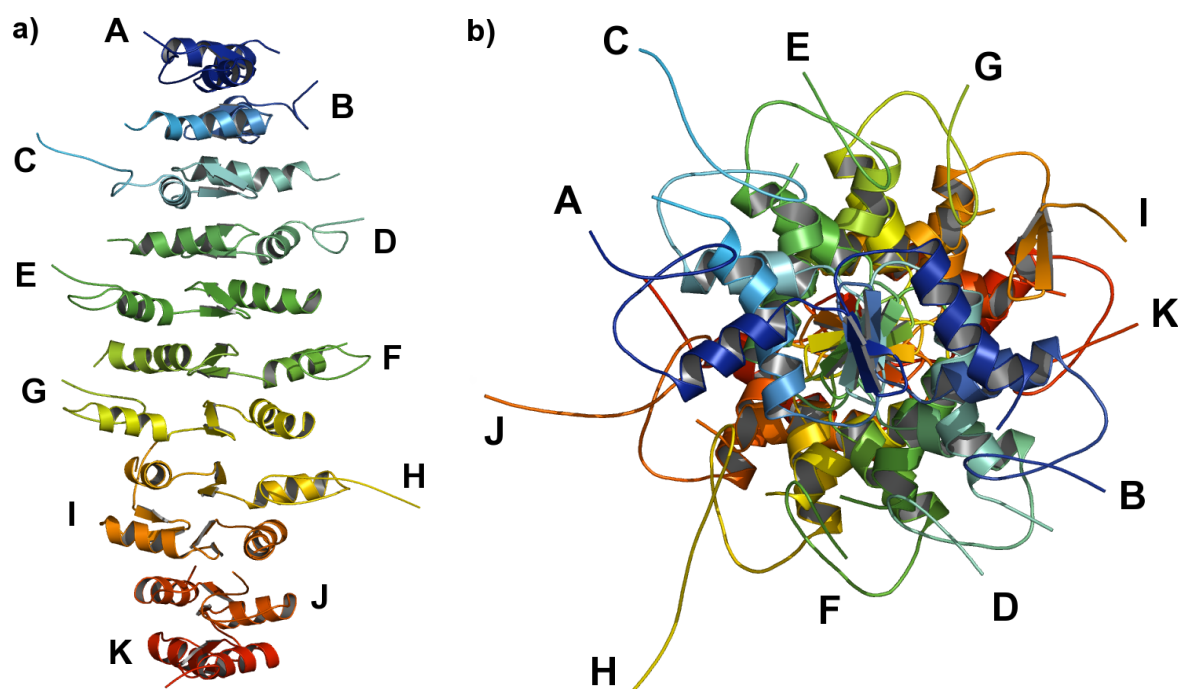


Figura 64. Empaquetamiento de las 11 cadenas de *UBZ1+2* en la unidad asimétrica, representadas de la *A* a la *K* y de distinto color. Se exponen dos orientaciones diferentes, **a)** a través del eje *C* de la celda, donde se observa la disposición paralela de una sobre otra, y **b)** a través del eje *A*, donde se muestra el giro en sentido horario de cada cadena con respecto a las colindantes. La letra que nombra cada cadena está situada en el extremo amino-terminal de la misma.

Los 22 átomos de zinc presentes en la unidad asimétrica, están coordinados por las cisteínas 10 y 13 y las histidinas 29 y 33 en el *UBZ1* y las cisteínas 37 y 40 y las histidinas 56 y 60 en el *UBZ2*. Los dos *UBZs* de los que consta cada cadena, mantienen una estructura muy similar, de forma que cada uno de ellos está formado una α -hélice, que ocupa aproximadamente un tercio de la secuencia, que está coordinado con el resto de la estructura por el átomo de zinc. En el caso del *UBZ2*, esta hélice está enfrentada a dos láminas β , como se observa en la Figura 65, donde se muestra la estructura de la cadena *C*, y la localización del átomo de zinc en cada uno de los dominios *UBZ*.

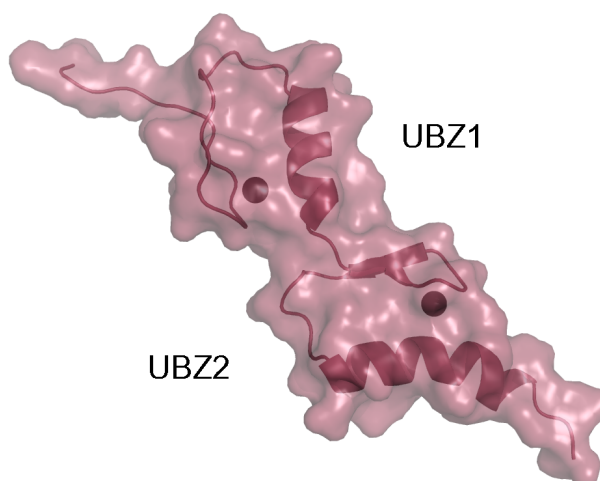


Figura 65. Diagrama de superficie mostrando la estructura de los dominios *UBZ1* y *UBZ2*, y la disposición del átomo de zinc (esferas) en cada dominio. La conformación en α -hélice se muestra de forma helicoidal, y las láminas- β como flechas.

Esta estructura responde a una coordinación en “dedo de zinc” (o “zinc finger”) del tipo Cys₂-His₂, con unas distancias medias aproximadas entre los átomos de zinc y los átomos de azufre de las cisteínas (-SG) de 2,3 Å y entre los átomos de zinc y los de nitrógeno (-NE₂) de las histidinas de 1,97 Å. Una representación de esta coordinación se muestra en la Figura 66, correspondiente al *UBZI* de la cadena *C*, donde también se representa el mapa de densidad electrónica localizado en ese área a 1,5 σ de nivel.

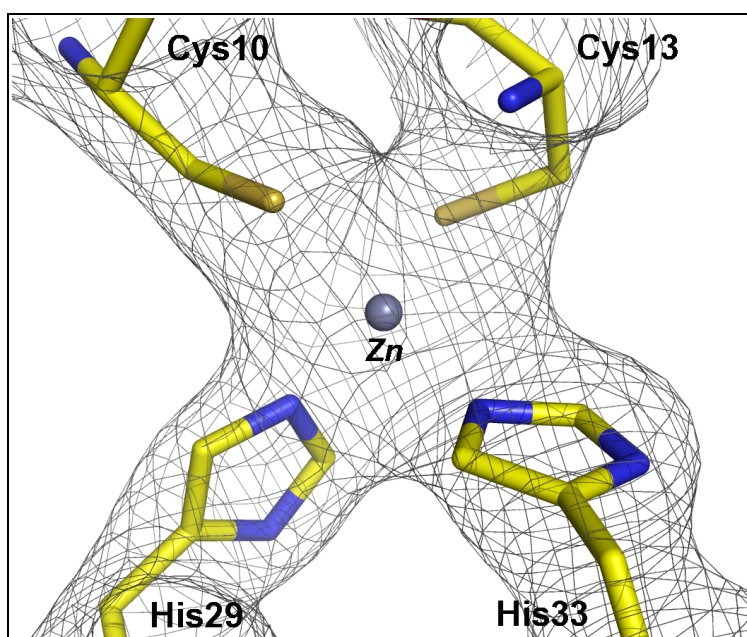
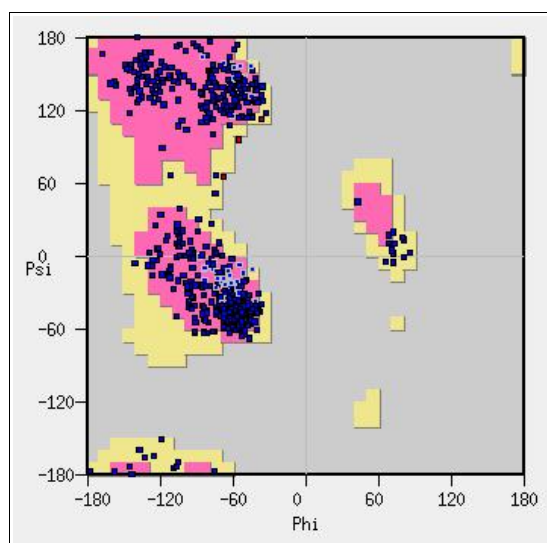


Figura 66. Coordinación entre Cys₂-His₂ y un átomo de zinc en el motivo “dedo de zinc” *UBZI*. Se muestra el mapa de densidad electrónica 2Fo - Fc a un nivel de 1,5 σ . La distancia entre el átomo de zinc y los azufres de las cisteínas 10 y 13 (color ámbar) es de 2,4 Å y 2,2 Å respectivamente y de 2 Å con los nitrógenos de las histidinas.

Con el fin de validar el modelo obtenido, se ejecutaron diversos programas. Uno de ellos fue la herramienta *Validate* del programa *Coot*, que se usó para calcular un diagrama de *Ramachandran* y comprobar así los residuos que podían presentar conformaciones desfavorables dadas por los ángulos diedros con los residuos colindantes. En este diagrama, mostrado en la Figura 67, se diferencian 594 residuos (el 90,69 % del total) en regiones preferidas (rosas en la figura), 58 aminoácidos (8,85 %) en regiones permitidas (amarillas) y 3 (0,46 %) en la región energéticamente desfavorable. Estos tres residuos están localizados en los extremos terminales de las cadenas.

Figura 67. Diagrama de Ramachandran conteniendo los 655 residuos de la estructura de UBZ1+2 que forman ángulos diedros con los colindantes. En rosa se representan las regiones favorables preferidas, en amarillo, las favorables permitidas y en gris el área donde se localizan los residuos energéticamente desfavorables.



Los parámetros que determinan el ajuste del modelo, estimados con herramientas como *Molprobability* y *PROCHECK* de *CCP4*, se muestran en la Tabla cristalográfica XX. Como mencionamos anteriormente, el *factor R* final resultó de 21,5 %, y el R_{free} de 25,5 %. Los *factores B* atómicos fueron de 26,9 Å² para la proteína, y de 20,25 Å² para los heteroátomos, lo que evidencia que la vibración del átomo con respecto a su posición central responde a unos valores estándares, sin ser muy elevados (aproximadamente 20 Å²). El parámetro de error o desviación cuadrática media estimada sobre las longitudes de los enlaces y de los ángulos fue de 0,01 Å y 1,349 ° respectivamente.

3. *Evidencias de interacciones de UBZ1+2 de Tax1BP1 con ubiquitina, poli-4Ub e Itch, por co-elución en una columna de exclusión de tamaño molecular.*

Una vez resulta la estructura de la región de la proteína TAX1BP1 que contiene ambos dominios potenciales de unión a ubiquitina, orientamos nuestro estudio a intentar obtener la estructura de dichas regiones en complejo con otras proteínas con las que son sabido, interaccionan.

Para tal fin, seleccionamos dos clases de proteínas, ambas relacionadas con la posible función que Tax1BP1 pudiera desempeñar en procesos de ubiquitinación y con ello en la regulación de importantes rutas de transducción. Éstas son, moléculas de ubiquitina o cadenas de las mismas formando poli-ubiquitinas, con las que interaccionarían para añadirlas a otras proteínas participantes en las mismas vías de señalización (Iha *et al.*, 2008), y la ligasa E3 de ubiquitina, denominada *Itch*, con la que interaccionaría a través de los motivos *PPxY* (Shembade *et al.*, 2008), siguiendo el mismo patrón de reconocimiento descrito (Sudol *et al.*, 1995), donde dominios *WW* (ricos en residuos de triptófanos, en este caso de *Itch*) reconocen dominios consenso *PPxY* (de Tax1BP1).

La formación de dichos complejos, tanto de la proteína *UBZI+2* con ubiquitina y poliubiquitina, como con los dominios *WW* de Itch, fueron analizados experimentalmente para ser aislados y llevados a ensayos de cristalización. Para tal fin, seleccionamos la aproximación de purificar cada componente de forma individual para ser posteriormente incubados e inyectados en una columna de exclusión por tamaño molecular, tanto las subunidades por separado, como la mezcla, lo que nos permitiría comparar sus comportamientos.

3.1 Interacción entre *UBZI+2* y ubiquitina

Para comprobar esta interacción, usamos ambas proteínas en solución, la proteína *UBZI+2* purificada, como ya hemos descrito previamente, y ubiquitina comercial (*BostonBiochem*[®]) disuelta en un tampón 20 mM TRIS pH 7,5 y 0,2 M NaCl, fueron incubadas a una concentración de 100 μ M durante 30 minutos a temperatura ambiente. Posteriormente, la mezcla fue concentrada para ser inyectada en una columna exclusión por tamaño molecular (“Superdex 75 16/60”), equilibrada con el mismo tampón.

La curva de absorbancia a una longitud de onda de 280 nm (Figura 68), mostró dos máximos bien diferenciados, correspondientes a unos volúmenes de elución de 72,9 mL y 82,5 mL respectivamente.

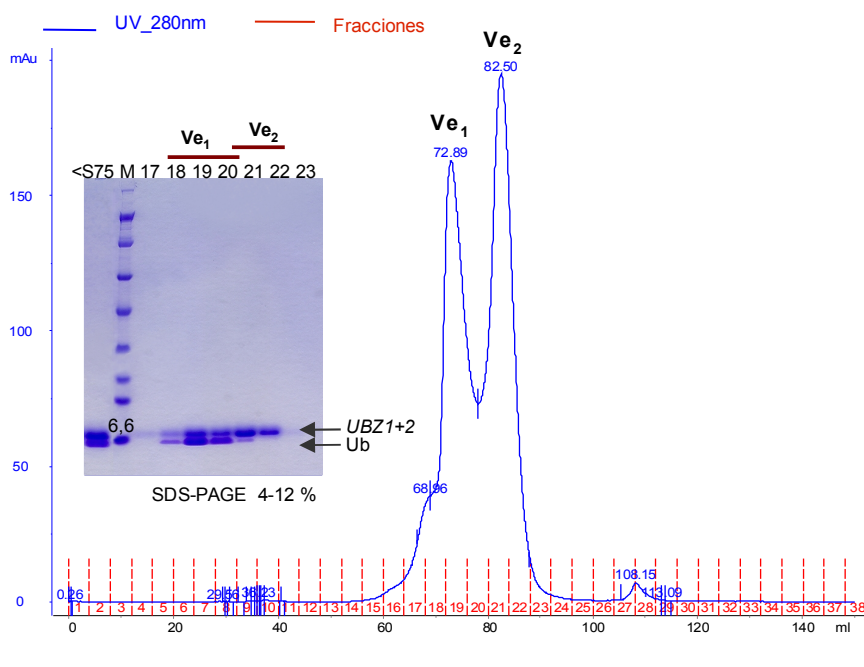


Figura 68. Cromatografía de exclusión por tamaño molecular de la mezcla proteica *UBZI+2* - Ubiquitina (*Ub*), analizadas mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 4-12 % BIS-TRIS (*Invitrogen*[®]). En el cromatograma se muestra la absorbancia a una longitud de onda de 280 nm, de las proteínas eluidas de una columna “Superdex 75 16/60”, en distintos volúmenes según su tamaño, y recogidas en fracciones de 4 mL, representadas en rojo. El tampón usado durante la elución contiene 20 mM TRIS pH 7,5 y 0,2 M NaCl.

Se muestra sobre cada pico de elución (“ Ve_1 ” y “ Ve_2 ”) su correspondiente volumen en mL. Las muestras antes de la inyección en la columna (<S75) y eluidas en los dos picos (18-20 y 21-22) se muestran en el gel de poliacrilamida. “M”: marcador de alto peso molecular, donde se representa el tamaño en kDa del peso cercano a las bandas de interés, 8,32 kDa para *UBZI+2* y 8,5 kDa para *Ub*.

Las fracciones correspondientes al primer volumen de elución (Ve_1 = de 72,9 mL y fracciones 18-20) muestran la presencia de las dos proteínas; *UBZI+2* y ubiquitina (*Ub*),

como se expone en el gel de poliacrilamida de la Figura 68. La interpolación de este volumen en una curva de calibración, se correspondió con un tamaño de aproximadamente 19 kDa.

El segundo volumen de elución (V_{e2} = de 82,5 mL y fracciones 21 y 22) mostró únicamente la presencia de la proteína *UBZ1+2*, con un peso estimado de 10 kDa.

El análisis del comportamiento de ambas proteínas por separado, muestra que eluyen en unos volúmenes diferentes a V_{e1} (donde co-eluyen al ser inyectadas en presencia de la otra). En conjunto, estos resultados confirman que bajo estas condiciones, existe interacción entre ellas, y además presumiblemente es lo bastante afin como para poder ser estudiada a través de filtración en gel y a bajas concentraciones.

Se ha descrito la existencia de diversas macromoléculas formadas por ubiquitina, implicadas en degradación y proteólisis. La formación de estas macromoléculas de cadenas de ubiquitinas (*poli-Ub*) está mediada a través de sus lisinas 48 o 63 (Pickart y Fushman, 2004). Recientemente, se ha demostrado mediante experimentos de inmunoprecipitación, que al menos el dominio *UBZ2* de Tax1BP1 interacciona también con diversas cadenas de ubiquitinas (del tipo 3-Ub, 4-Ub, 5-Ub o 6-Ub), donde la interacción mas fuerte es la mantenida con 4-Ub (Iha *et al.*, 2008). Por este motivo decidimos estudiar la interacción entre el fragmento que contiene los dos *UBZs* con una cadena formada por cuatro ubiquitinas, para posteriormente intentar su cristalización.

Las cuatro ubiquitinas (*poli-4Ub*) fueron expresadas en tándem en un sistema procariota, mediante el vector pGEX4T1-4Ub, facilitado por el Dr. Dikic. Para ello, dicho vector fue transformado en células *Rosetta(DE3)* con el fin de ser expresado durante 6 horas, bajo una concentración de 1 mM de agente inductor IPTG y a 37 °C. Tras obtener el cultivo celular, se procedió a la lisis bacteriana y centrifugación para extraer la fracción proteica soluble, que posteriormente fue sometida a purificación, según el protocolo expuesto en la sección “3.2” de *Materiales y Métodos*.

El proceso de purificación se desarrolló en diferentes etapas. En primer lugar, se inyectó la fracción soluble del lisado celular resultante de la sobre-expresión de *GST-poli-4Ub*, en una columna de afinidad por GST y fue eluida con 10 mM de glutatión reducido. La proteína de fusión se sometió a proteólisis con trombina para eliminar la cola de GST. Tanto la proteína no digerida como la cola de GST liberada se eliminaron mediante la utilización de una segunda columna de afinidad por GST. Por último, la fracción no retenida de esta columna fue concentrada e inyectada en una columna de filtración en gel, del tipo “Superdex 75 26/60”, previamente equilibrada con el tampón 20 mM TRIS pH 7,5 y 50 mM NaCl. Finalmente, la proteína *poli-4Ub* eluyó en un volumen de 156 mL, que tras analizar mediante un patrón de calibración previamente establecido, se obtuvo un peso molecular aproximado de 36,6 kDa. Teniendo en cuenta que una molécula de *poli-4Ub* son aproximadamente 35 kDa, los datos sugieren que la proteína tendría una disposición oligomérica de monómero (los datos en cuanto a expresión y purificación de esta proteína no se muestran).

La proteína *poli-4Ub* final, que presentó un elevado grado de pureza, fue incubada con *UBZ1+2* a temperatura ambiente, durante 30 minutos y posteriormente se analizó el patrón de elución de la mezcla en la misma columna de exclusión por tamaño molecular que fue usada en las purificaciones de las dos proteínas individuales. Esta cromatografía, expuesta en la Figura 69, mostró tres máximos de absorbancia de UV a 280 nm. El primero se correspondía con un volumen de 126,5 mL (V_{e1} = fracciones 26-28). Tras

calcular el peso molecular correspondiente a este volumen, en una curva de calibración con estándares conocidos, se sugirió un tamaño teórico de 72 kDa. Los otros dos máximos, a unos volúmenes de 156,5 mL (Ve_2 ; fracciones 32 a 35) y 211,1 mL (Ve_3 ; fracciones 42-44) correspondían a unos pesos moleculares de 34,8 kDa y 9,34 kDa respectivamente.

El análisis de estas fracciones mediante electroforesis en gel de poliacrilamida, mostró la presencia de ambas proteínas en el primer máximo de elución (Ve_1), de *poli-4Ub* en el segundo (Ve_2) y de *UBZI+2* en el tercero (Ve_3).

Estos resultados muestran que existe un claro desplazamiento y co-elución de ambas proteínas *poli-4Ub* y *UBZI+2* en el pico Ve_1 , con respecto al comportamiento que muestran tras ser inyectadas de forma individual en la misma columna. Con ello, sugerimos que, bajo estas condiciones existe una evidente interacción entre los dos motivos “dedos de zinc” *UBZI+2* de Tax1BP1 con las cuatro moléculas de ubiquitina. Las dos moléculas, que de forma individual eluyen con posterioridad en los dos máximos minoritarios, se corresponden con el comportamiento que mantienen cuando son inyectadas por separado.

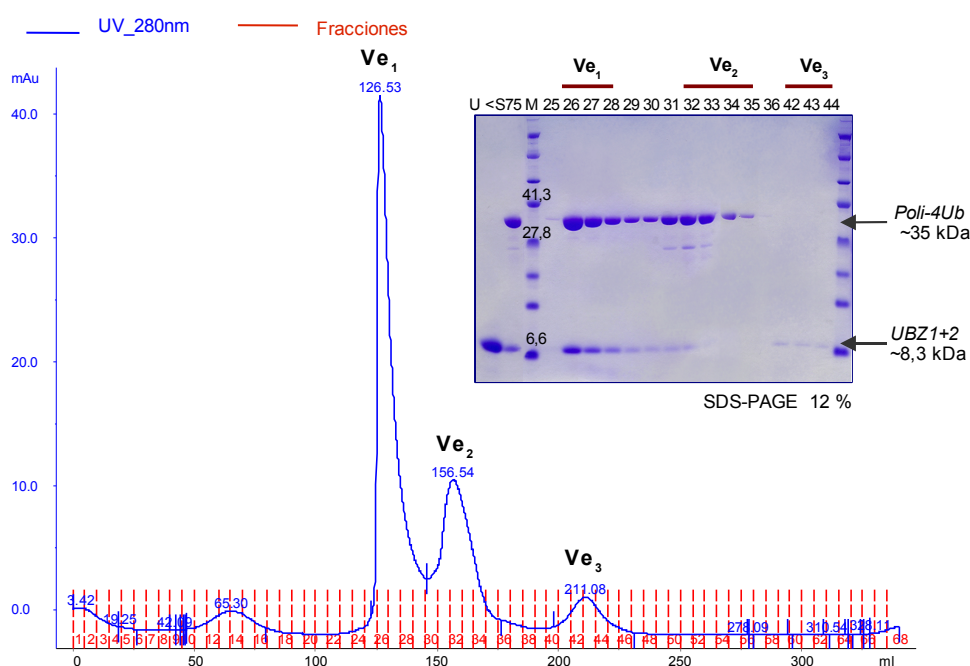


Figura 69. Cromatografía de exclusión por tamaño molecular de la mezcla proteica formada por *UBZI+2* y por *poli-4Ub*, analizadas mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 12 % BIS-TRIS (*Invitrogen*^l). En el cromatograma se muestra la absorbancia a una longitud de onda de 280 nm, de las proteínas eluidas en distintos volúmenes según su tamaño y recogidas en fracciones de 5 mL, representadas en rojo. Se expone sobre cada pico de elución (Ve_1 , Ve_2 y Ve_3) su correspondiente volumen en mL. Las muestras *UBZI+2* pura (U), la mezcla proteica antes de la inyección en la columna (<S75) y las eluidas en los tres picos (26-28, 32-35 y 42-44) se muestran en el gel de poliacrilamida. “M”: marcador de alto peso molecular, donde se representa el tamaño en kDa de los peso cercanos a las bandas de interés.

3.2 Interacción entre UBZ1+2 e Itch

Por otro lado, y como hemos mencionado anteriormente, también se ha descrito recientemente la interacción entre la ligasa de ubiquitina Itch, y UBZ1+2 de Tax1BP1, mediante ensayos de co-inmunoprecipitación (Shembade *et al.*, 2008). Estos estudios muestran que la unión entre ambas proteínas es mediada por la secuencia conservada PPxY de Tax1BP1, localizada dentro de sus dos dominios UBZs (Figura 5). Esta interacción es inhibida cuando la tirosina de dicho dominio es mutada. Por otro lado, Itch contiene cuatro dominios ricos en pares de triptófanos (WW), descritos como motivos de unión a dominios PPxY (Sudol *et al.*, 1995) (Figura 70-A). Por todo ello, decidimos analizar la interacción entre Itch y Tax1BP1, concretamente entre sus dominios WW y PPxY.

Para ello diseñamos tres construcciones que comprendían diferentes regiones de la proteína Itch, la primera abarcando los cuatro motivos WWs (WW₁-WW₄), la segunda los dos primeros dominios (WW₁+WW₂) y la última los dos restantes (WW₃+WW₄). Estos tres fragmentos se muestran con detalle en la Figura 70-B.

Las secuencias de ADN humanas codificante para las tres construcciones fueron clonadas, tal y como se describe en el apartado “I” de *Materiales y Métodos*, en un vector de expresión procariota (pGEX6P2) para ser expresadas como proteína de fusión a GST en su extremo amino-terminal. Dichas proteínas de fusión presentan un sitio de corte para la proteasa PreScission, que facilita la eliminación del fragmento GST, dejando 5 residuos extras (GPLGS) en su extremo amino-terminal.

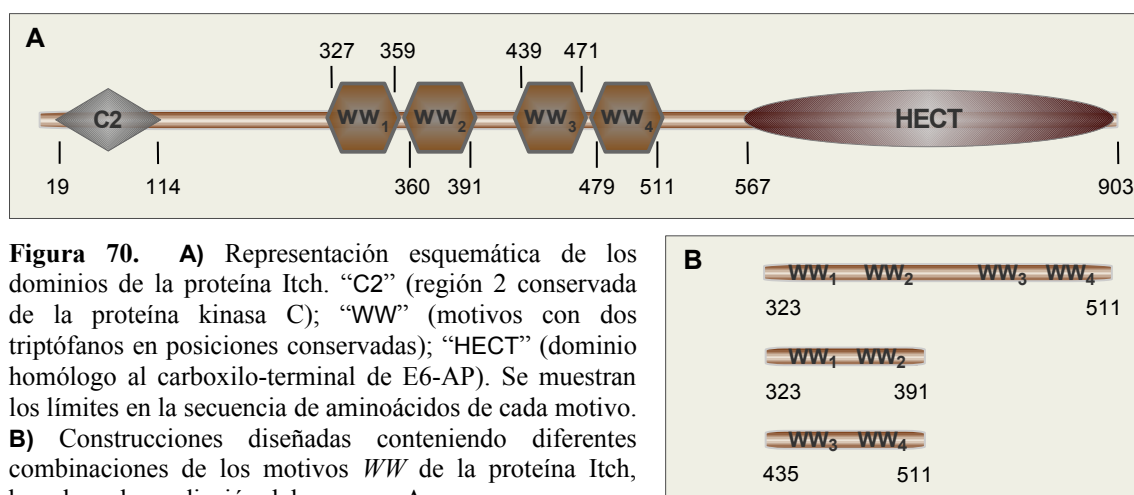


Figura 70. **A)** Representación esquemática de los dominios de la proteína Itch. “C2” (región 2 conservada de la proteína quinasa C); “WW” (motivos con dos triptófanos en posiciones conservadas); “HECT” (dominio homólogo al carboxilo-terminal de E6-AP). Se muestran los límites en la secuencia de aminoácidos de cada motivo. **B)** Construcciones diseñadas conteniendo diferentes combinaciones de los motivos WW de la proteína Itch, basada en la predicción del esquema A.

Tras diferentes ensayos de expresión, que se llevaron a cabo de la forma descrita en *Materiales y Métodos*, se optimizaron las condiciones de sobre-expresión de las tres construcciones. Dichas condiciones englobaban la utilización de la cepa *Rosetta(DE3)* como célula de expresión, la inducción de expresión con 0,5 mM de IPTG durante unas 15 horas y a una temperatura de 15 °C. Pese a los numerosos ensayos de optimización realizados tras la lisis bacteriana del cultivo celular, con el fin de favorecer la solubilidad de las proteínas, las construcciones WW₁-WW₄ y WW₁+WW₂ se obtuvieron en la fracción insoluble. Por el contrario, la construcción soluble WW₃+WW₄, de la que se obtuvieron los

niveles de sobre-expresión expuestos en la Figura 71-A, pudo ser sometida a las diferentes etapas de purificación, tal y como se detalla en el apartado 3.2 de *Materiales y Métodos* para purificar proteínas fusionadas a GST.

En primer lugar, WW_3+WW_4 fue inyectada en una columna de afinidad por GST, donde eluyó tras aplicar 10 mM de glutatión reducido. Las muestras que contenían GST-PP- WW_3+WW_4 fueron sometidas a proteólisis con el fin de eliminar la cola de GST para, tras ser inyectada en una columna de filtración en gel (“Superdex 75 26/60”), separar la proteína WW_3+WW_4 libre de GST. En la cromatografía de exclusión por tamaño molecular llevada a cabo a 4 °C de temperatura y en un tampón con 20 mM TRIS pH 7,5 y 0,2 M NaCl, obtuvimos cinco máximos de absorbancia a 280 nm (Figura 71-B). Tras analizar estas fracciones mediante electroforesis en gel de poliacrilamida, se comprobaron que los máximos referidos a los volúmenes de 147,6 mL (fracciones 27 a 32) y 218,6 mL (fracciones 42 a 49), correspondían a la cola de GST y a la proteína de interés WW_3+WW_4 digerida, respectivamente. Dichos volúmenes de elución fueron interpolados en una curva de calibración, sugiriendo un peso molecular aproximado para el segundo volumen de 8 kDa (WW_3+WW_4), y una conformación monomérica, si tenemos en cuenta que dicha proteína tiene un tamaño de 9,5 kDa.

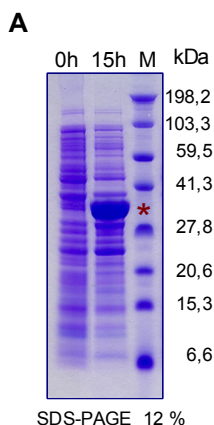
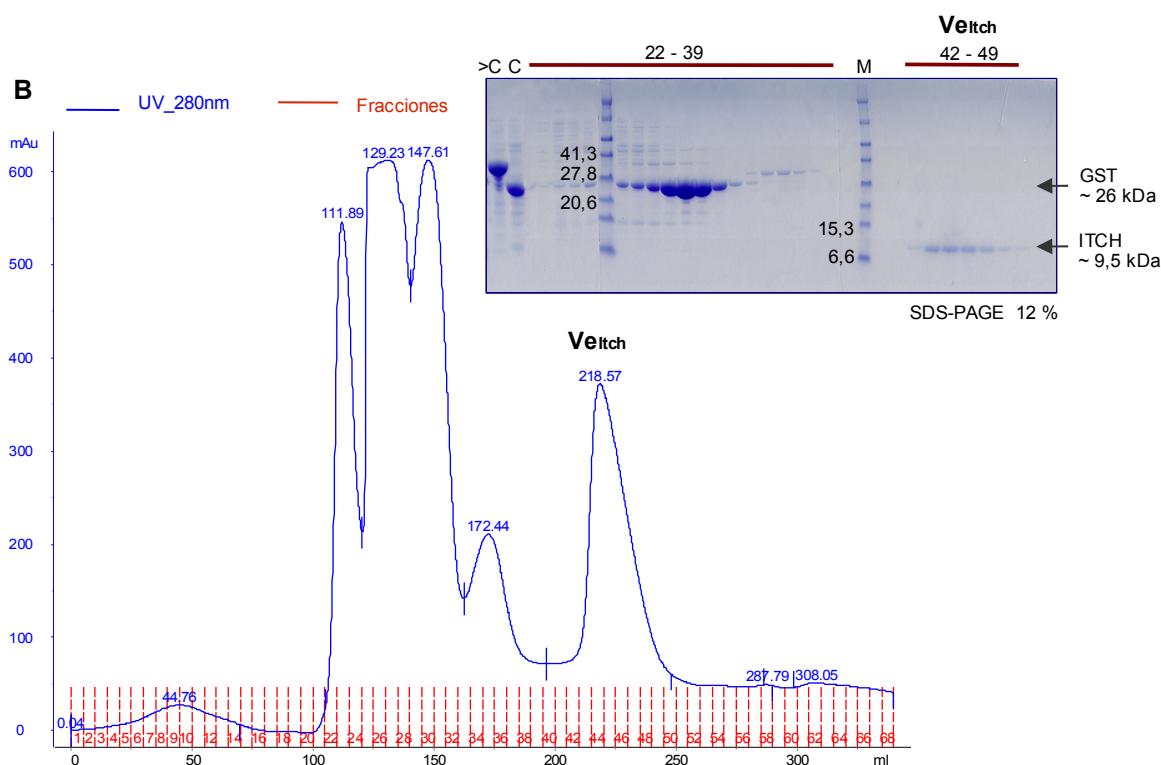


Figura 71. A) Expresión de GST-PP-*Itch* $_{WW_3+WW_4}$, en cepa de *E.coli*, analizada mediante electroforesis en gel de poliacrilamida (SDS-PAGE) al 12 %, BIS-TRIS.

El carril marcado como “0h” se corresponde con el total de proteínas expresadas antes de inducir la sobre-expresión de la proteína heteróloga mediante adición de IPTG. El carril “15h” contiene el total de proteínas presente en las células transcurridas 15 horas de añadir el agente inductor.

Para referenciar los tamaños moleculares se usó un marcador (M) de alto peso molecular (*BioRad*^E) cuyos pesos moleculares se muestran a la derecha en unidades de kDa. El asterisco rojo indica la banda correspondiente a la proteína de fusión de interés sobre-expresada, de aproximadamente 36 kDa.



B) Muestras resultantes de la proteólisis de GST-PP-*Itch* y cromatografía de exclusión por tamaño molecular de la proteína digerida, analizadas mediante electroforesis en gel de poliacrilamida (SDS-PAGE) del 12 % BIS-TRIS (*Invitrogen*[®]).

En el cromatograma se muestra la absorbancia a una longitud de onda de 280 nm, de las proteínas eluidas de una columna “Superdex 75 26/60” en distintos volúmenes según su tamaño y recogidas en fracciones de 5 mL, en rojo. Se expone sobre cada pico de elución, su correspondiente volumen, siendo “*Veitch*” (218,57 mL) el de la proteína de interés, que corresponde a las fracciones 42 a la 49, y el volumen de 147,6 mL (fracciones 27 a 32) a la cola de GST eliminada.

En el gel de poliacrilamida, los carriles “<C” y “C” se corresponden con muestras tomadas de la mezcla proteica antes y después de la proteólisis respectivamente. “M”: marcador de alto peso molecular, donde se representa el tamaño en kDa de los pesos cercanos a las bandas de interés.

Tras concentrar las fracciones que mostraban la presencia de *WW*₃+*WW*₄ (42 a 49) hasta una concentración de 100 μM, la muestra resultante fue incubada con la proteína *UBZ1*+2 purificada, a una proporción equimolar. Esta mezcla fue concentrada hasta un volumen de 4,5 mL para ser inyectada en la misma columna de exclusión por tamaño molecular que las usadas para las purificaciones de ambas proteínas individuales. Tras eluir las proteínas según su tamaño, en un tampón conteniendo 20 mM TRIS pH 7,5 y 0,2 M NaCl, se obtuvieron dos máximos bien definidos (Figura 72). El primero de ellos, correspondía a un volumen de 180,3 mL y las fracciones 36 a 38, mientras que el segundo (201,2 mL) fue recogido en las fracciones 40 a 46. Con el fin de identificar las proteínas contenidas en dichas fracciones, se llevó a cabo una electroforesis SDS-PAGE. Tras su visualización con azul de “Coomassie” comprobamos que el primer máximo (180,3 mL) contenía ambas proteínas, *UBZ1*+2 y *WW*₃+*WW*₄, como se puede observar en el gel expuesto en la Figura 72, mientras que en las fracciones del volumen de elución de 201,2 mL sólo estaba presente la proteína individual *WW*₃+*WW*₄. Ambos máximos de absorbancia corresponden a unos pesos moleculares interpolados en una curva de calibración de 19,64 kDa y 11,8 kDa respectivamente.

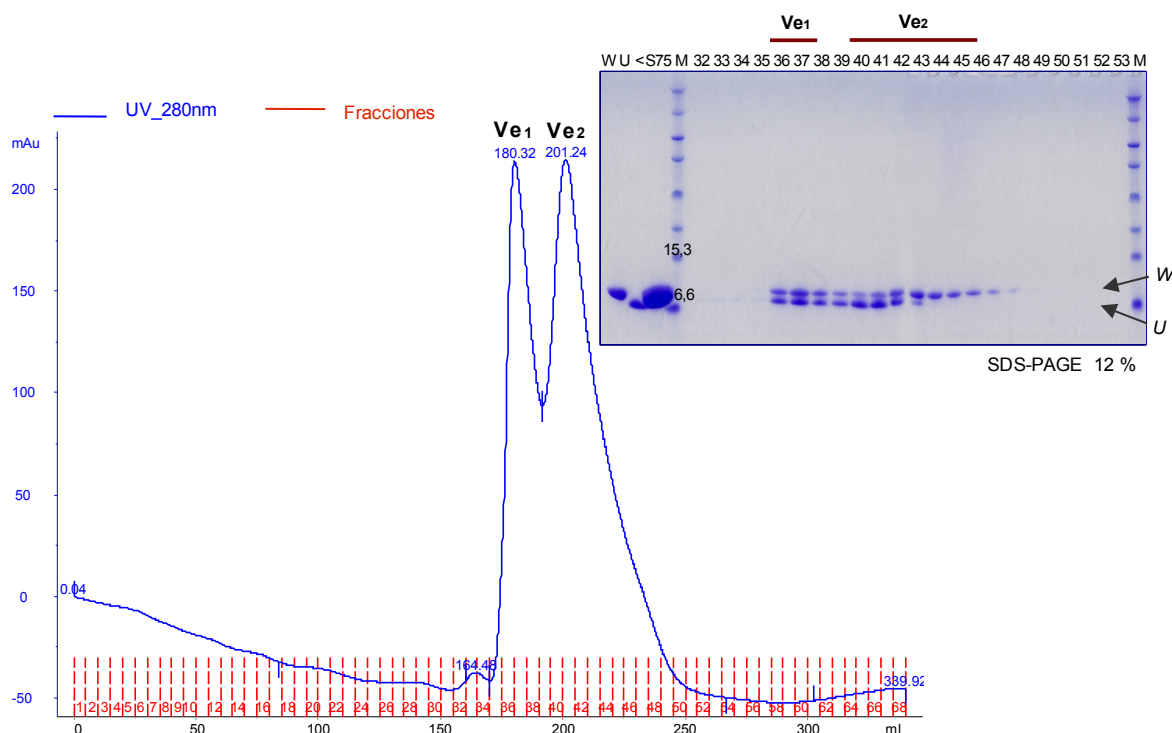


Figura 72. Cromatografía de exclusión por tamaño molecular de la mezcla proteica conteniendo los dominios *UBZ1+2* de Tax1BP1 y *WW₃+WW₄* de Itch y electroforesis en gel de poliacrilamida de las muestras eluidas en la columna. La curva de absorbancia a una longitud de onda de 280 nm muestra dos máximos (“Ve₁” y “Ve₂”), correspondientes a las proteínas eluidas de la columna “Superdex 75 26/60” según su tamaño molecular, y visualizadas en el gel SDS-PAGE al 12 % BIS-TRIS (*Invitrogen*[®]) teñido con solución de azul de “Coomassie”. En el gel se exponen con números las fracciones de 4 mL colectadas (representadas en rojo en el cromatograma), así como las englobadas dentro de cada máximo de absorbancia. “W” hace referencia a los motivos *WW₃+WW₄* de Itch, con un peso molecular de 9,5 kDa, mientras que “U” alude a los dominios *UBZ1+2* de Tax1BP1, de aproximadamente 8,3 kDa. El marcador usado (M) fue de alto peso molecular.

Estos resultados sugieren que bajo estas condiciones, ambas proteínas interactúan, co-eluyendo en un volumen diferente al que lo hacen de forma individual. Además, tal y como se observa tras comparar los pesos moleculares obtenidos, esta interacción tendría lugar en una proporción equimolar, uniéndose un monómero de *UBZ1+2* con otro de *WW₃+WW₄*, ya que el peso molecular estimado para ese máximo (19,6 kDa) se aproxima a la suma de ambos, respectivamente 8,3 kDa y 9,5 kDa.

4. Ensayos de cristalización de *UBZ1+2* en complejo con Ub, poli-4Ub e *ItchWW₃+WW₄*

Tras obtener evidencias de las interacciones entre *UBZ1+2* de Tax1BP1 y la ubiquitina, la poli-4Ub y los dominios *WWs* de Itch, nos centramos en intentar cristalizar dichos complejos. Para tal fin, las fracciones resultantes de las cromatografías de exclusión por tamaño molecular que contenían las dos proteínas, es decir; las fracciones

18 a 20 en el caso de mono-ubiquitina, las 26 a 28 en el caso de poli-4Ub y las 36 a 38 para Itch, fueron concentradas hasta 60, 35 y 55 mg/mL respectivamente.

Con las muestras obtenidas tras el proceso de concentración, se probaron alrededor de 700 condiciones distintas de cristalización, repitiendo estos experimentos a diferentes temperaturas (4 °C, 16 °C y 22 °C), concentraciones de proteína y proporciones entre proteína:reservorio, y utilizando la técnica de difusión de vapor en gota sentada.

Finalmente, surgieron cristales de las muestras *UBZI+2 + Ub* y de *UBZI+2 + poli-4Ub* en gotas de 1,8 µL de volumen, con una proporción equivolumétrica entre proteína:reservorio y a una temperatura de incubación de 22 °C. La mezcla que presentaba la presencia de *UBZI+2* y mono-ubiquitina cristalizó en dos condiciones diferentes: 0,1 M HEPES pH 7,2; 1,5 M Li₂SO₄ y 0,1 M CHES pH 9,5; 0,2 M Li₂SO₄; 1 M K/Na-Tartrato. Los cristales, mostrados en la Figura 73-A, surgieron tras 3 días de comenzar el experimento, y a una concentración de 20 mg/mL. Para el caso de *UBZI+2 + poli-4Ub*, los cristales (Figura 73-B) aparecieron transcurridos 5 días de mezclar la proteína, a una concentración de 60 mg/mL, con la condición de cristalización conteniendo 0,1 M TRIS pH 8,5 y 0,75 M MgCl₂. Estos cristales se reprodujeron con igual forma y tamaño en condiciones conteniendo 0,8 M Acetato de Magnesio ó 0,7 M Formato de Magnesio, lo que nos sugirió que el Magnesio era esencial para cristalizar la proteína. Por ello, la proteína fue incubada con diferentes concentraciones de sales de magnesio antes de realizar ensayos de cristalización. Finalmente, no se consiguieron optimizar estos cristales siguiendo dicha aproximación.

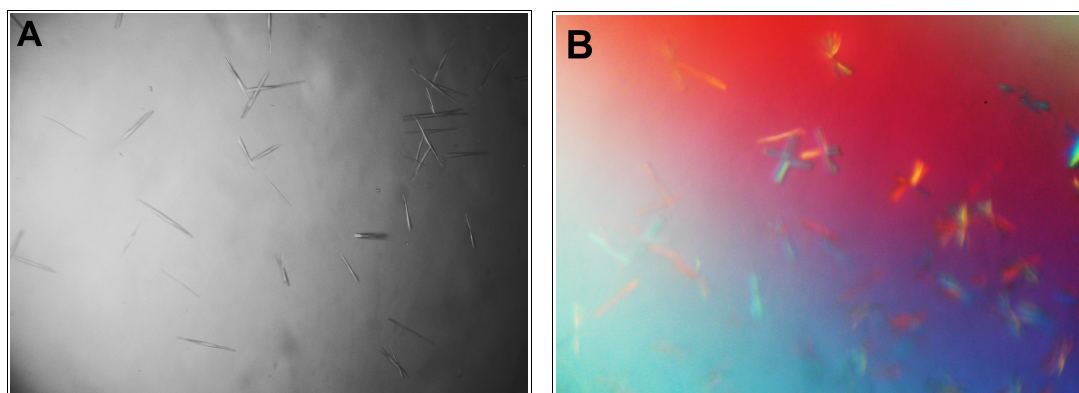


Figura 73. Cristales de los dominios *UBZI+2* en presencia de ubiquitina (**A**) y de *poli-4Ub* (**B**), surgidos mediante difusión de vapor, en gota sentada de 1,8 µL de volumen y una proporción 1:1 (v/v) entre proteína:reservorio, y bajo una temperatura de 22 °C. Las condiciones de cristalización contienen respectivamente 0,1 M HEPES pH 7,2 y 1,5 M Li₂SO₄ para los cristales del panel A, y 0,1 M TRIS pH 8,5 y 0,75 M MgCl₂, para los del B.

Ambos tipos de estructuras cristalinas fueron irradiados, sin éxito, bajo un haz de rayos X en el sincrotrón de Grenoble. Con el uso de muy diversas técnicas, expuestas en el apartado “8.2” de la sección de *Materiales y Métodos*, se intentaron optimizar los cristales surgidos en ambas condiciones, con el fin de obtenerlos únicos y sobretodo, de mayor tamaño. Finalmente, no logramos mejorar los cristales, y con ello, tampoco la difracción.

Discusión

La metástasis representa uno de los procesos menos conocidos, abarcando mecanismos que hacen de la enfermedad una de las principales causas de muerte hoy en día. La mayor esperanza para cualquier investigador implicado en metástasis es que la comprensión de los mecanismos de supresión de la misma, desemboque en el desarrollo de nuevas terapias. Algunos de los avances más importantes que se han llevado a cabo contra el cáncer, en las últimas décadas, ha sido en el diseño de nuevas drogas. Éstas han surgido como consecuencia del esclarecimiento y comprensión de la biología molecular de la enfermedad. Por ello, es indispensable profundizar en el conocimiento a nivel molecular de todas las rutas que controlan el desarrollo metastásico, y de aquellas que se encuentran des-reguladas durante su progresión.

Si estudiamos las vías de transducción implicadas, tanto en la regulación positiva como negativa de la metástasis, observamos que la vía de activación del factor de transcripción NFκB es una de ellas, involucrada principalmente en la regulación de la expresión de genes implicados en invasión, metástasis, angiogénesis (Karashima *et al.*, 2003) y apoptosis celular (Barkett y Gilmore, 1999; Gapuzan *et al.*, 2002; Karin, 2006; Lin y Karin, 2003; Sliva *et al.*, 2002), y cuya activación constitutiva se relaciona con diversos tipos de cáncer. Por ello, la mejor manera de estudiar su regulación sería adentrándonos en el conocimiento de antagonistas de la ruta. Es aquí donde entran en juego diferentes proteínas conocidas como supresores de metástasis (Stafford *et al.*, 2008; Steeg, 2003), que según un grupo de investigadores, forman parte de una firma intrínseca basada en un patrón de expresión diferencial a los que rigen el tumor primario (Kang *et al.*, 2003; Minn *et al.*, 2005; Smid *et al.*, 2006). Una de dichas proteínas relacionada además con la regulación negativa de NFκB es BRMS1, que inhibe metástasis de mama, melanoma y ovario principalmente (Seraj *et al.*, 2000b).

Si además analizamos los mecanismos de regulación relacionados con dicha ruta, observamos la importancia de los procesos de ubiquitinación en el control de la actividad de diferentes antagonistas de la vía, proceso también relacionado estrechamente con cáncer (Haglund y Dikic, 2005; Hoeller *et al.*, 2006). Es aquí donde cobra importancia otro efector negativo (al igual que BRMS1) de esta vía de señalización. Se trata de Tax1BP1 (Iha *et al.*, 2008), proteína de unión a la oncoproteína Tax1, un potente activador de la ruta (Ballard *et al.*, 1988; Leung y Nabel, 1988). La proteína Tax1BP1 forma parte de un complejo multiproteico que regula la vía del factor de transcripción mediante procesos de ubiquitinación (Shembade *et al.*, 2008). Los mecanismos de ubiquitinación relacionados con el ciclo celular, han sido ampliamente estudiados en el laboratorio con anterioridad, lo que también nos condujo a la elección de esta proteína como objeto de estudio.

Es preciso resaltar el hecho de que quizás las mejores dianas terapéuticas no sean los supresores de metástasis o las proteínas codificadas por estos, sino algunos de sus ligandos, bien agonistas o antagonistas. Por tanto, el conocimiento de las rutas en las que las proteínas BRMS1 y Tax1BP1 están implicadas y de sus ligandos, representa uno de los puntos de mayor interés para el presente estudio. Una de la mejores aproximaciones para alcanzar este objetivo es el conocimiento de las estructuras tridimensionales de ambas proteínas o de dominios relevantes involucrados en sus funciones, además de sus complejos, lo que aportaría información para ayudar en el esclarecimiento de los mecanismos moleculares que acontecen en el desarrollo de la enfermedad. La determinación de la estructura cristalina de nuevas proteínas va generando una sustancial

base de datos que correlaciona estructura/función, por lo que parece cada vez más plausible que el conocimiento de la estructura se pueda utilizar como herramienta primaria para identificar actividad biológica (Banfield *et al.*, 1998).

Es importante mencionar que hasta el momento no existe ningún estudio estructural de ninguna de las proteínas objeto de estudio, lo que caracteriza al presente trabajo de novedoso. A ello se suma la dificultad que implica adentrarnos en el estudio de proteínas de las que no se conoce nada a nivel estructural, y muy poco a nivel funcional.

La técnica elegida para abordar los estudios estructurales ha sido la Cristalografía de Rayos X, la cual es una técnica compleja que precisa de unos estrictos requerimientos en cuanto a pureza química y estructural de la muestra para la obtención de cristales.

La utilización de sistemas heterólogos para la sobre-expresión de proteína recombinante ha supuesto además otras ventajas adicionales como son la incorporación de etiquetas que facilitan los posteriores pasos de purificación. Entre los inconvenientes de la expresión en sistemas heterólogos procariotas, se encuentran formaciones de cuerpos de inclusión. Una aproximación alternativa para sobreponerse a este inconveniente es el establecimiento de protocolos de repliegamiento, sin embargo creímos interesante insistir más en el diseño de construcciones múltiples para cada gen, que aumentasen la probabilidad de obtener construcciones que resultaran en la producción de proteína soluble. Esta solución pareció más eficaz que el establecimiento de protocolos de repliegamiento, como está quedando de manifiesto en los grandes proyectos de genómica estructural existentes hoy en día en Europa y Estados Unidos principalmente.

Debido a la ausencia de información estructural para llevar a cabo un diseño racional de construcciones génicas, se ha utilizado en el presente trabajo toda la información disponible a partir de la secuencia de los genes de interés, como pueden ser los alineamientos múltiples, la comparación de dominios secuenciales y los elementos de conservación evolutiva, como se puede observar en el apartado de *Resultados*.

1. Estudios estructurales de BRMS1 y su interacción con NMI

El gen *brms1* codifica para una proteína de 246 aminoácidos, de la cual existen relativamente pocos datos de funcionalidad y nulos en cuanto a su estructura. La predicción de dominios de la proteína BRMS1 expuesta en la Figura 3, muestra la existencia de una región amino-terminal rica en residuos de glutámicos, otras dos centrales ricas en leucinas y dos posibles señales de localización nuclear en el extremo carboxilo. Los análisis realizados con diferentes programas de predicción de motivos “coiled-coil” mostraron la existencia de dos de estos motivos, comprendidos entre los aminoácidos 51-78 y 147-179 (Figura 9). Estos motivos “coiled-coil” desempeñan un papel muy importante en interacciones entre proteínas y en funciones de gran relevancia biológica (Lupas, 1996; Mason y Arndt, 2004). En particular, BRMS1 interacciona con diferentes proteínas a través de sus motivos “coiled-coil”, tal es el caso de la unión con la proteína ARID4A, que forma parte del complejo Sin3-HDAC de remodelado de la cromatina, para cuya interacción parece ser suficiente el segundo de los motivos (Hurst *et al.*, 2008).

1.1 Caracterización de la interacción entre BRMS1 y NMI

Otros ejemplos de interacciones de BRMS1 a través de uno de sus motivos “coiled-coil”, concretamente el primero, han sido demostradas previamente en el laboratorio por el Dr. José Rivera mediante estudios de doble híbrido en levadura y confirmadas posteriormente mediante diversos ensayos (artículo en revisión). La enorme importancia de algunas de las proteínas humanas rescatadas en este ensayo, hizo centrar aún mas nuestra atención en este primer motivo “coiled-coil” de BRMS1. Una de ellas fue NMI, un co-factor de la transcripción que interacciona con diversos transductores y activadores de este proceso, como N-myc, c-Myc o proteínas de la familia STAT, modulando importantes vías de señalización (Bannasch *et al.*, 1999; Zhu *et al.*, 1999). Si observamos la Figura 10, NMI consta, entre otros, de un motivo predicho en “coiled-coil” entre los residuos 21 a 64 (Chen *et al.*, 2000). Si a ello añadimos que una región mínima de NMI rescatada en los estudios anteriormente referidos que interacciona con el primer “coiled-coil” de BRMS1 engloba también el motivo “coiled-coil”, además del dominio *IFP35*, incrementa aún mas nuestras sospechas de que la interacción entre ambas proteínas pudiera estar mediada por dichos motivos.

En el presente trabajo, decidimos centrar nuestro interés en el estudio estructural del primer “coiled-coil” de BRMS1 individual y en complejo con las regiones “coiled-coil” e *IFP35* de NMI, para lo que diseñamos numerosas construcciones (Tabla VII). Algunas de ellas fueron diseñadas teniendo en cuenta los límites predichos de ambos motivos (51 a 78 para BRMS1 y 21 a 64 para NMI), para aumentar la probabilidad de limitar el motivo estructural real. De esta forma, se añadió y eliminó un total de 5 residuos a cada uno de los extremos, generando 4 construcciones mas para cada proteína. La sobre-expresión en sistema heterólogo procariota nos permitió obtener seis de los ocho fragmentos solubles, confirmando la utilidad de este tipo de aproximación que ha sentado sus bases en el éxito de los grandes proyectos de genómica estructural establecidos en varios países. En el caso de NMI, los fragmentos 21-64 y 21-69 no se consiguieron sobre-expresar en sistema procariota, a diferencia de las construcciones 16-64 y 16-69, que además de obtener elevados niveles de proteína expresada, resultaron soluble el primero e insoluble el segundo, lo que implica que tan sólo 5 residuos de diferencia son capaces de causar un comportamiento muy distinto.

A pesar de que todas las proteínas y polipéptidos solubles fueron destinados a ensayos de cristalización, dos de ellas (BRMS₅₁₋₈₄ y NMI₁₆₋₆₄) fueron caracterizadas biofísicamente, así como su interacción proteína-proteína. Los estudios biofísicos mediante espectroscopía, sugirieron que ambos polipéptidos mantenían un promedio de estructura secundaria típico de α -hélice. Como pudimos observar en las respectivas cromatografías de exclusión por tamaño molecular, las conformaciones oligoméricas mas probables de dichos polipéptidos en solución son de dímero para BRMS₅₁₋₈₄ y de trímero para NMI₁₆₋₆₄.

La utilización de técnicas tales como resonancia magnética nuclear (RMN) y microcalorimetría (ITC), nos permitió comprobar la interacción entre ambas construcciones bajo las condiciones descritas. Además, los resultados obtenidos en la técnica de RMN sugirió que esta interacción es muy específica, ya que tenía lugar incluso disminuyendo la concentración de ambos componentes hasta 30 μ M. De igual forma, pudimos obtener los parámetros estequiométricos, que se ajustaba a un modelo de dos

sitios de unión, con unas afinidades del orden de micromolar (0,55 μ M y 2,5 μ M para los sitios N1 y N2 respectivamente).

Los resultados obtenidos en el espectro bidimensional de RMN, muestran una conformación desestructurada del motivo “coiled-coil” de BRMS1 de forma individual, ya que no se disciernen señales definidas dadas por sus protones y heteronúcleos. Sin embargo, los experimentos de Dicroísmo Circular (CD) sugieren una estructura secundaria en α -hélice con dos mínimos claros de absorbancia en 208 y 222 nm. Esta controversia se puede explicar teniendo en cuenta que RMN se trata de una técnica mas sensible que la de CD, y que los valores estudiados en cuanto a la última, se traducen en un promedio de la estructura en general presente en la muestra. En último lugar, podríamos sugerir que esta diferencia de resultados es debida a la presencia de distintos componentes en la composición de los tampones usados (10 mM HEPES pH 7,5 y 50 mM NaCl para CD y 20 mM Fosfato Sódico pH 6,5; 50 mM NaCl y 1 mM DTT para RMN), aunque parece difícil pensar que ésta sea la causa de una diferencia tan grande.

Por otro lado, también se analizó el comportamiento de las proteínas completas BRMS1 y NMI. Los resultados obtenidos en estos ensayos, expuestos en el apartado “6.1” de *Resultados*, basados en la expresión simultánea (co-expresión) de los componentes participantes, sugieren la existencia de interacción entre las dos proteínas. A través de este método obtuvimos el complejo formado por ambas proteínas, purificándolo mediante cromatografía de afinidad por histidinas (mantenida únicamente por uno de los componentes). Con esta muestra, conseguimos cristales de la mezcla proteica (Figura 52), que aunque no se pudieron optimizar para conseguir un patrón de difracción que permitiese la resolución de la estructura, representan un indicio que apoya la posibilidad de insistir en esta técnica para obtener información estructural de las proteínas completas.

De todas las construcciones llevadas a ensayos de cristalización, obtuvimos cristales de la construcción BRMS1₅₁₋₈₄, y tras numerosos intentos por asignar las fases y resolver así la estructura, conseguimos obtener la conformación tridimensional de dicho motivo “coiled-coil” predicho, usando para ello la señal anómala y las diferencias isomorfas originadas por mercurio, mediante *SIRAS*.

1.2 Estructura del primer motivo “coiled-coil” de BRMS1

La obtención de cristales del polipéptido BRMS1₅₁₋₈₄ nos permitió conseguir un patrón de difracción adecuado y posteriormente un modelo tridimensional del mismo, que corroboró la predicción de estructura en motivo “coiled-coil” de dicho fragmento. El modelo fue validado, obteniéndose unos *factores R* y *R_{free}* de 20,5 % y 26,8 % respectivamente, que marcan las diferencias relativas entre los factores de estructura observados (*F_{obs}*) y calculados (*F_{calc}*). La fiabilidad del modelo trazado se consideró elevada considerando que los parámetros con respecto a la geometría y ajuste de densidad electrónica (Figura 33) fueron adecuados, y que el diagrama de *Ramachandran* (Figura 32) solamente mostró un residuo (E9 de la cadena G) en la región no permitida, el cual se trata de un aminoácido terminal.

El modelo obtenido (Figura 28), mostró la presencia de 8 cadenas helicoidales en la unidad asimétrica dispuestas en pares unidos por puentes disulfuro. Estos resultados evidencian que el ambiente sin agente reductor favorece el estado oxidado de la proteína,

la formación de estos enlaces y con ello la aparición de los cristales. Además, cuando los ensayos de cristalización fueron realizados usando un tampón que contenía 2 mM DTT, los cristales tardaban mas de un mes en aparecer, lo que supone multiplicado por cuatro el tiempo en el que aparecían en el tampón con 20 mM TRIS pH 7,5 y 50 mM NaCl, sin agente reductor.

De un total de 280 residuos, se pudieron modelar un gran número de ellos (208), tal y como se expone en la Tabla XIV. El empaquetamiento de las cadenas en el cristal y la diferencia entre las superficies de contacto entre ellas, fueron las causas por las cuales no se obtuvo densidad electrónica que pudiese ser usada para construir los residuos situados en los extremos de las hélices. Si analizamos el diagrama mostrado en la Figura 74, podemos especular sobre la causa en la diversidad obtenida en las longitudes de las cadenas. Posiblemente, al incrementar el número de contactos que existen entre ellas se generen estructuras mas estables; es decir, con átomos menos vibrantes alrededor de su posición central (Figura 74-A). Además, el incremento del *factor B* es mas pronunciado en los extremos de todas las cadenas, y sobretodo en aquellas que no están modeladas completamente. Este efecto, que ya se mostraba en la sección de *Resultados* (Figura 34), se observa también en la Figura 74. En ella se percibe cómo las cadenas que mantienen mas puntos de contacto con las adyacentes son las mas completas y además sus átomos tienen menor valor de vibración (*factor B*).

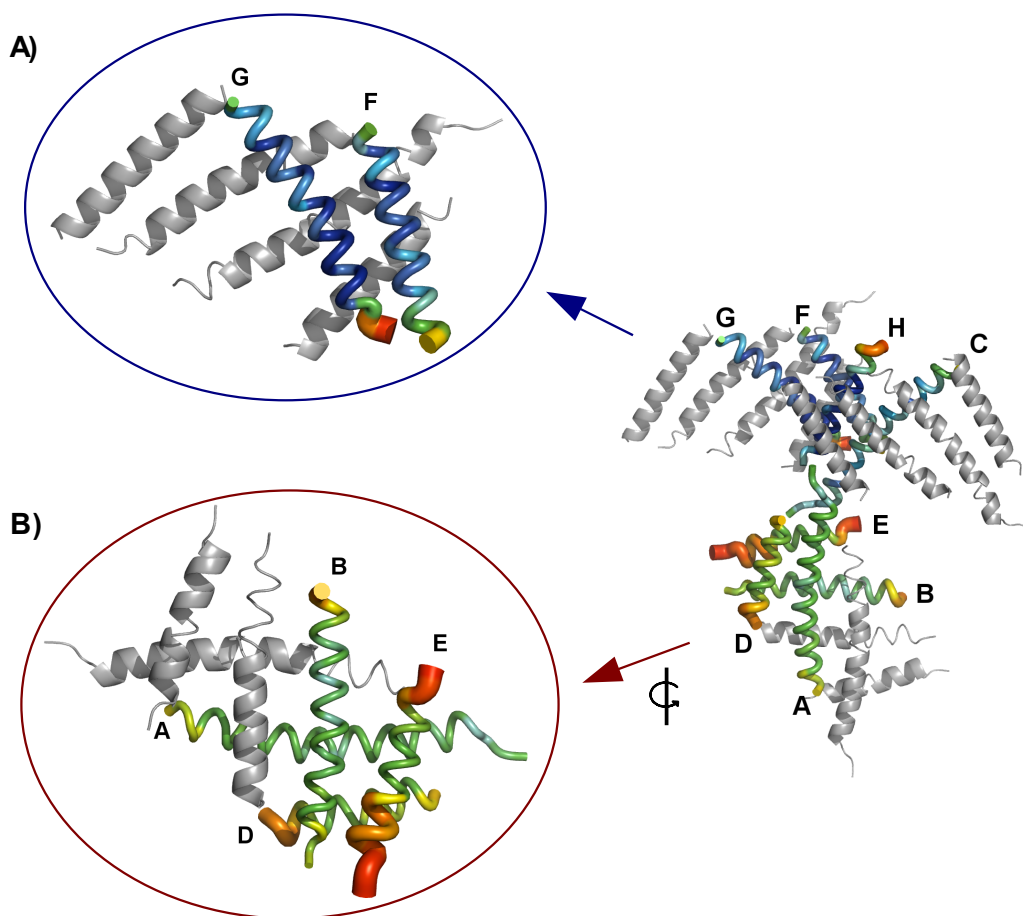


Figura 74. Cadenas de la unidad asimétrica, mostrando las diferencias en los *factores B* atómicos, y simétricas colindantes, del modelo estructural del polipéptido BRMS1₅₁₋₈₄. Los *factores B* se representan

según una escala de colores, donde el color azul se refiere a los átomos con menor valor, que aumenta sucesivamente hasta alcanzar el color rojo. A la derecha se muestra de forma general, las cadenas nombradas de la unidad asimétrica, y en color gris algunas simétricas con las que mantienen contactos. En los paneles **A)** y **B)** se exponen la diversidad de longitudes trazadas para diferentes cadenas, según los puntos de contactos que las estabilizan. Se muestran, como ejemplos, cadenas con menos vibración de sus átomos (*F* y *G*), con 3 y 4 contactos, en el panel **A)**. En el panel **B)** se expone la diferencia entre cadenas mas elongadas (*A*), con cuatro puntos de contacto, y otras mas cortas con dos contactos (*D* y *E*).

De este modo, el número de puntos de contacto que mantiene cada cadena con las adyacentes, es proporcional a su longitud: cuatro puntos de apoyo, para moléculas con 35 residuos (*A* y *C*), tres para una longitud de 27 aminoácidos (ej. cadena *B*), y dos para las mas cortas, como la cadena *D* con 18 residuos, que exponen por lo tanto gran parte de sus residuos al solvente.

Analizando dicho empaquetamiento de todas las cadenas en el cristal se explica la dificultad que supuso obtener cristales del primer motivo “coiled-coil” predicho de la proteína BRMS1. El cambio experimentado por la muestra BRMS1₅₁₋₈₄ al añadirle el polipéptido NMI₁₆₋₆₄, generando señales definidas en el espectro bidimensional de RMN, evidencia que la interacción entre ambas proteínas dirige al complejo a un sistema estructurado. Observando todos los resultados globamente, sugerimos que para la formación de cristales es necesario bien, el establecimiento de puentes disulfuro entre cadenas de BRMS1₅₁₋₈₄, o bien la presencia de otra molécula con la que interacciona. Ambos casos estabilizan la estructura helicoidal favoreciendo el establecimiento de contactos entre las cadenas y, con ello, la aparición de cristales.

1.2.1 Twister, análisis de motivos “coiled-coil”

Una vez obtenida la estructura y con el fin de corroborar la predicción de este fragmento (BRMS1₅₁₋₈₄) de ser motivo “coiled-coil”, buscamos en la celdilla pares de moléculas que pudiesen responder a la geometría de estos motivos, tras lo que localizamos dos pares de cadenas que mantenían una morfología aparentemente típica de “coiled-coil”: *A-A_{sim2}* y *G-C_{sim5}* (donde *A_{sim2}* sería la cadena *A* del grupo simétrico 2 y *C_{sim5}*, la cadena *C* del grupo simétrico 5). Los análisis realizados sobre la cadena *A* con su simétrica y la *G* con la simétrica de *C*, demuestran que responden a una geometría de tipo “coiled-coil” antiparalelo y levógiro (Figura 31). Posteriormente, se delimitó la región de contacto entre las dos cadenas (del residuo 11 al 33) y fueron examinados con el programa *Twister*, el cual determina los parámetros estándares de “coiled-coil” como una función de número de residuos, en base a criterios estructurales (Strelkov y Burkhard, 2002).

Este programa asignó a la estructura primaria del polipéptido BRMS1₅₁₋₈₄, un patrón secuencial conforme a la “heptada” típica de estos motivos. Dicho patrón contiene siete residuos en tándem denominados *abcdefg* y repetido *n* veces, donde las posiciones *a* y *d* corresponden principalmente a residuos apolares, como Leu, Ile o Val, lo que crea condiciones óptimas en la superficie de estas cadenas para facilitar un empaquetamiento hidrofóbico (Strelkov y Burkhard, 2002). La correspondencia entre la secuencia en aminoácidos y la asignación del patrón en tándem, se expone en la Figura 75.

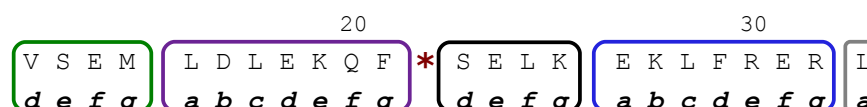


Figura 75. Asignación sobre la secuencia de BRMS1₅₁₋₈₄ del patrón de 7 residuos repetidos en tándem, o “heptada” típico de los motivos estructurados en “coiled-coil”. Las secuencias de la proteína y de la asignación del patrón *a-g*, se muestran en la primera y segunda fila respectivamente. Las “heptadas” se delimitan con recuadros, y se expone la correspondencia numérica de los residuos en la secuencia. Con un asterisco se localiza dónde se produce una interrupción o discontinuidad en el patrón. Para el análisis se usó el programa *Twister*.

En la estructura formada por el par de moléculas *A-A_{sim2}*, observamos que las interacciones entre ambas cadenas son principalmente hidrofóbicas, mantenidas entre los residuos Val11 de una cadena con Leu33 y Leu15 con Phe29 y Leu33 y viceversa. Además, estos residuos hidrofóbicos se localizan en posiciones *a* y *d* de la “heptada”, cumpliendo una característica típica de los motivos “coiled-coil”. Las distancias mantenidas entre las cadenas laterales de estos residuos hidrofóbicos son de entre 3,9 y 4,2 Å, que se encuentran dentro de rangos permitidos de enlaces hidrofóbicos (4,5 Å o mas próximo). Como se observa en la Figura 76, estos grupos de residuos, representados en rojo, establecen interacciones localizadas en ambos extremos de la doble hélice antiparalela, donde se originan superficies hidrofóbicas.

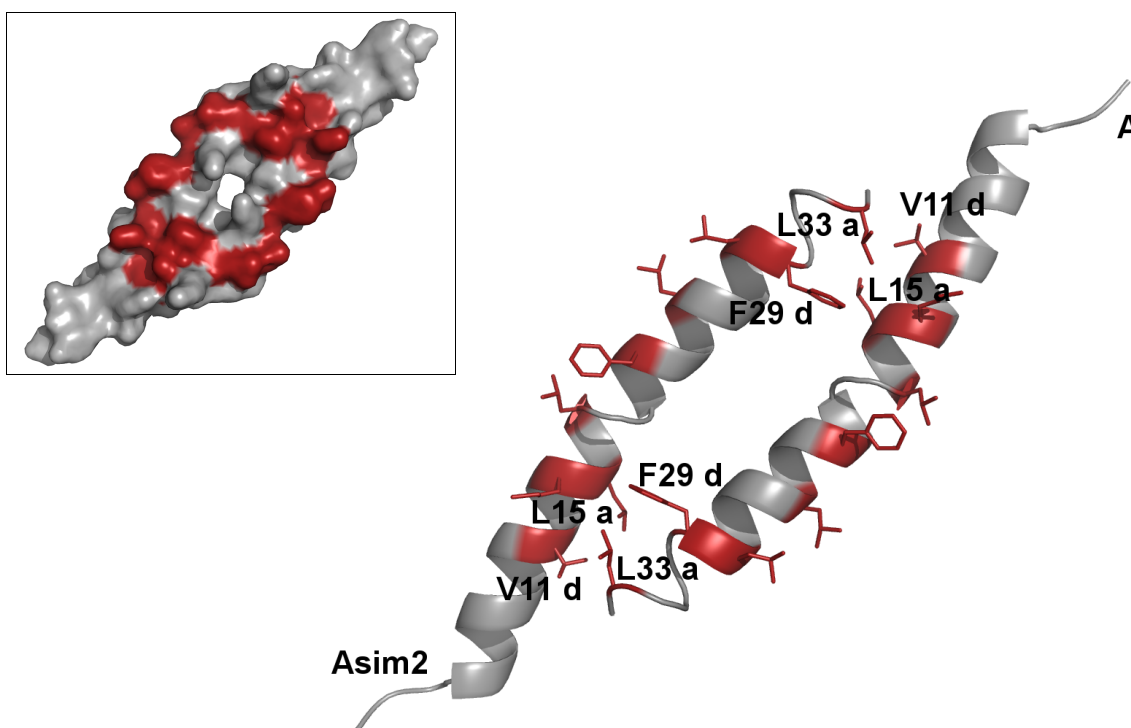


Figura 76. Superficie de hidrofobicidad e interacciones hidrofóbicas establecidas entre las cadenas *A-A_{sim2}* de la estructura de BRM1₅₁₋₈₄, a través de residuos hidrofóbicos localizados en las posiciones *a* y *d* de la “heptada”, manteniendo el patrón típico de motivos “coiled-coil”. Se representan las cadenas laterales y en color rojo todos los aminoácidos de carácter hidrofóbico de la estructura, e identificados con el número y la letra que ocupan en la secuencia y en la “heptada” respectivamente, aquéllos que están implicados en interacciones de este tipo entre ambas cadenas. Las cadenas están identificadas en sus extremos amino-terminales.

Además de estas interacciones hidrofóbicas, existen otras de carácter polar en la zona central de la superficie de contacto entre las dos cadenas, establecidas principalmente entre Glu26 de una cadena con Glu18 y Ser22 de la otra, residuos que también ocupan posiciones *a* y *d* en la “heptada” (Figura 77-A). Al tratarse de una doble cadena antiparalela, la disposición de estos grupos de residuos interaccionando entre sí se repite de la misma manera dos veces.

Como se muestra en la Figura 75, en la secuencia aminoacídica se localiza tras el residuo 21 una discontinuidad o interrupción del patrón de la “heptada” (en inglés “stutter”), típico en estos motivos, como demuestran algunos artículos que lo corroboran (Brown *et al.*, 1996; Lupas, 1996). Sin embargo, una vez resuelta la estructura, pudimos contemplar que esta discontinuidad en la secuencia no distorsiona la morfología de la hélice. Por otro lado, existe una interrupción helicoidal observada en todas las cadenas, tras la Leu17, que abarca el Glu18 y la Lys19, tal y como muestra la Figura 77-A.

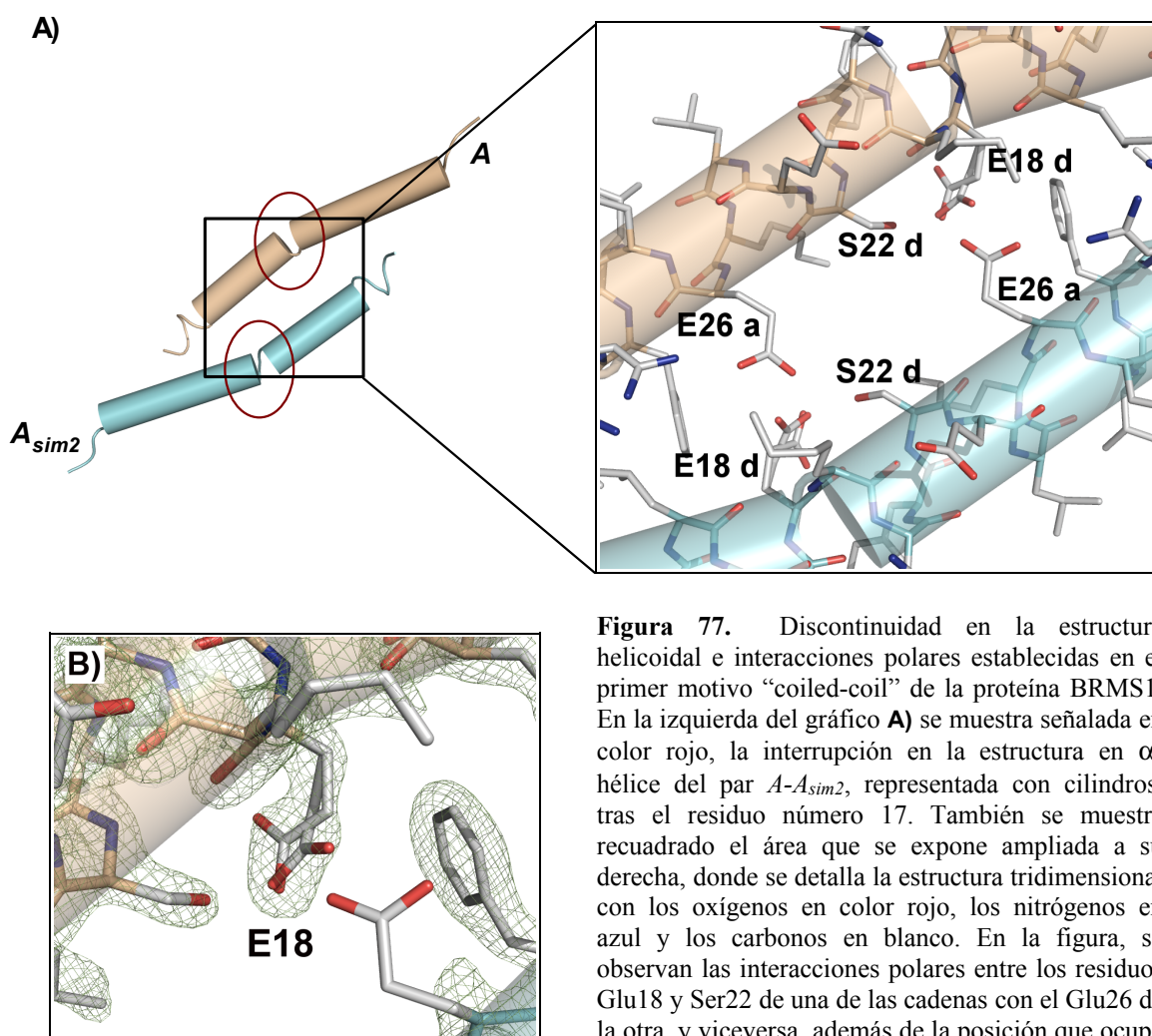


Figura 77. Discontinuidad en la estructura helicoidal e interacciones polares establecidas en el primer motivo “coiled-coil” de la proteína BRMS1. En la izquierda del gráfico **A)** se muestra señalada en color rojo, la interrupción en la estructura en α -hélice del par A - A_{sim2} , representada con cilindros, tras el residuo número 17. También se muestra recuadrado el área que se expone ampliada a su derecha, donde se detalla la estructura tridimensional con los oxígenos en color rojo, los nitrógenos en azul y los carbonos en blanco. En la figura, se observan las interacciones polares entre los residuos Glu18 y Ser22 de una de las cadenas con el Glu26 de la otra, y viceversa, además de la posición que ocupa cada uno en la “heptada”.

B) Detalle de mapa de densidad electrónica 2Fo - Fc trazado a un nivel de 1,5 σ , en torno al modelo construido de ambas cadenas. Se identifica el residuo E18 de la cadena *A*, cuya cadena lateral se expone doble.

La discontinuidad helicoidal observada en la estructura puede ser debida a la diferencia del ángulo diedro entre los residuos 18 y 19, que aunque se localizan en la región preferida en el diagrama de *Ramachandran* (Figura 32), presenta una conformación mas abierta que el resto de los ángulos de la hélice. Las cadenas laterales de algunos residuos Glu18 fueron modeladas con una doble conformación y con un 50 % de ocupación cada una. Esta dos alternativas se observaron claramente en el mapa de densidad electrónica obtenido, mostrado en la Figura 77-B, y podrían estar originadas por el carácter alternativo presentado en la interacción del residuo con el Glu26 y la Ser22.

Si además comparamos las orientaciones de cada residuo de la “heptada” en la estructura con las típicas de los motivos “coiled-coil”, mostradas en la Figura 78-A (<http://cis.poly.edu/~jps/coilcoil.html>), observamos que mantienen exactamente el mismo patrón, orientándose las cadenas laterales de los residuos localizados en las posiciones *a* y *d* hacia el interior de la doble hélice. Los residuos restantes (*bcefg*) exponen sus cadenas laterales hacia el exterior, colocándose en el mismo área aquellos con la misma correspondencia en la “heptada”. De esta forma, los residuos *a* y *d* son los que mantienen la doble estructura, creando un ambiente hidrófobo que facilita su empaquetamiento (Figura 78-B).

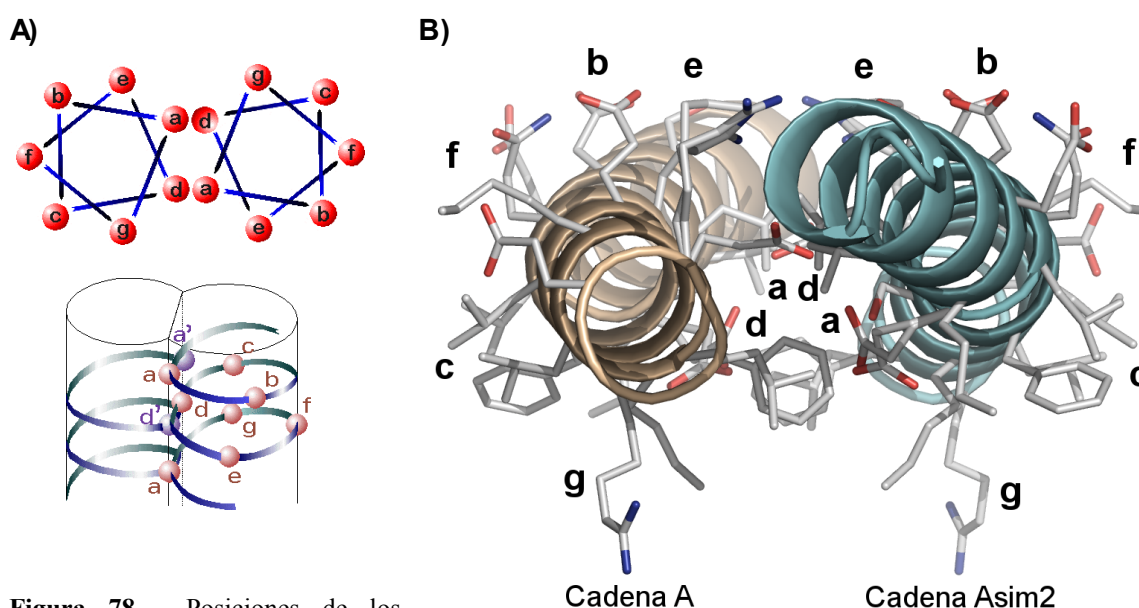


Figura 78. Posiciones de los residuos y orientación de sus cadenas laterales en las dos hélices antiparalelas *A-A_{sim2}* formando el “coiled-coil”. **A)** Posiciones de los residuos de la “heptada” típicas de motivos “coiled-coil”. En este caso el diagrama corresponde a una doble hélice paralela, observada a través de su eje longitudinal (arriba) y transversal (abajo). Figura adaptada de <http://cis.poly.edu/~jps/coilcoil.html>. **B)** Vista a través del eje longitudinal de la doble hélice *A-A_{sim2}* de la estructura del polipéptido BRMS151-84, donde se exponen los grupos de residuos con su correspondencia en la “heptada”. Los aminoácidos con la misma asignación se localizan en la misma zona, por lo que se identifican con una única letra.

La estructura adoptada por la doble hélice del polipéptido BRMS151-84, que responde a los patrones típicos de una conformación en “coiled-coil”, evidencia que realmente se trata de un motivo de este tipo, y que posiblemente está desempeñando un importante papel como mediador de interacciones entre proteínas. Si analizamos la predicción de motivos de la proteína (Figura 3), observamos que carece de otros motivos estructurales definidos. Solamente la región rica en residuos de glutámicos presente en el

extremo amino-terminal de la proteína podría estar involucrada, *a priori*, en interacción con otras proteínas y factores de transcripción. Tal y como está descrito, muchas proteínas median su interacción con factores de transcripción a través de regiones ricas en glutámicos.

BRMS1 contiene además una región rica en leucinas (L-x₆-L-x₆-L-x₆-L), ocupando los aminoácidos 67 a 88, dentro de este primer motivo “coiled-coil”. Este patrón recuerda a motivos estructurales de la familia de proteínas *bZIP*, que consiste en 53 factores de transcripción que utilizan el “coiled-coil” rico en residuos de leucinas básico como motivo de dimerización (Newman y Keating, 2003; Vinson *et al.*, 2006). Aunque no existen estudios que encuadran a la proteína BRMS1 dentro de esta familia, unos resultados obtenidos recientemente pueden estar estrechamente relacionados con su función. En dichos ensayos evalúan la especificidad en la dimerización y para ello diseñan proteínas que contienen el motivo “coiled-coil” rico en leucinas de las proteínas *bZIP*, además de una extensión acídica anfipática en el extremo amino-terminal, que también está presente en BRMS1 (región rica en glutámicos). Estas proteínas, denominadas como *aZIP*, forman heterodímeros con los dominios *bZIP* de los factores de transcripción a través de la región acídica, inhibiendo su unión al ADN (Acharya *et al.*, 2006). Si analizamos la presencia de motivos “coiled-coil” tanto en BRMS1, que contiene también la región acídica, como en NMI, catalogado como co-factor de la transcripción, podemos vaticinar que quizás la interacción entre ambas proteínas sucede de manera dependiente de residuos de leucina, al igual que en el caso *bZIP-aZIP*.

Tras resolver la estructura de esta construcción (aminoácidos 51 a 84), nos planteamos la opción de clonar un fragmento elongado en su extremo carboxilo-terminal con el fin de visualizar el giro de la doble hélice que se predice para completar el motivo “coiled-coil” y poder seguir analizando el patrón de la “heptada”. Este fragmento, del aminoácido 51 a 98, comprende 15 residuos mas, que serían suficientes para analizar la continuidad de la doble hélice tridimensional y del patrón de la “heptada” en la estructura primaria, tal y como se observa en un modelo confeccionado en la Figura 79, donde se ha continuado con un fragmento de un tamaño aproximado de 15 aminoácidos por el extremo carboxilo-terminal.

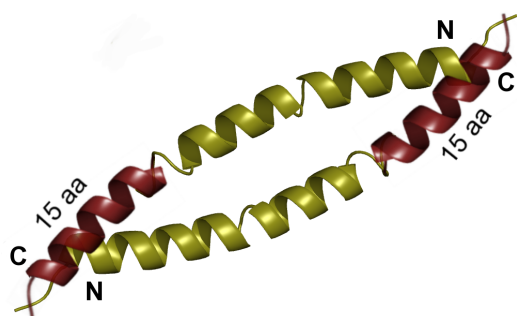


Figura 79. Modelo de estructura elongando 15 residuos en el extremo carboxilo-terminal del motivo “coiled-coil” de BRMS₅₁₋₈₄, con el fin de simular la continuación de la doble hélice. El fragmento cuya estructura se ha resuelto se muestra de color verde y la extensión de los 15 residuos en color rojo. Las abreviaturas “N” y “C” aluden a los extremos amino- y carboxilo-terminales respectivamente, y “aa” a aminoácidos.

Analizamos si existía continuidad del patrón de “heptada” en el extremo elongado, continuando con el ya asignado en la secuencia de la construcción BRMS₁₅₁₋₈₄. Observamos que en tres de las cuatro posiciones *a* y *d* extras se localizan residuos hidrofóbicos, y además dos de ellos son leucinas (Figura 80), evidencias que apuntan que esta extensión aminoacídica seguiría formando parte del motivo “coiled-coil”.

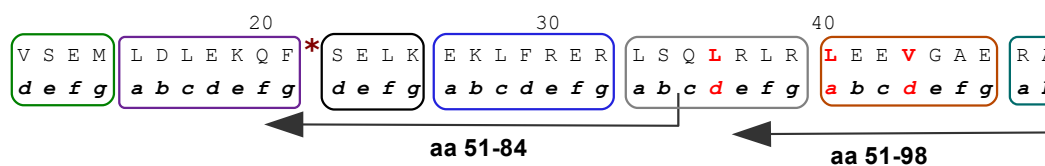


Figura 80. Asignación del patrón de 7 residuos repetidos en tándem o “heptada”, sobre la secuencia aminoacídica de la construcción clonada para BRMS1₅₁₋₉₈, versión elongada que engloba al primer motivo “coiled-coil”, cuya estructura tridimensional se ha resuelto. En la primera fila se muestra la secuencia primaria y en la segunda la correspondencia con el código *a-g* de la “heptada”, que se delimitan con recuadros. Las posiciones *a* y *d* extras en la construcción BRMS1₅₁₋₉₈, donde se localizan residuos hidrofóbicos están representadas en color rojo. “Heptada” asignada con el programa *Twister*. Con un asterisco se localiza dónde se produce una interrupción o discontinuidad en el patrón.

Esta construcción fue clonada, expresada y purificada con éxito. En un intento de resolver las fases, perseguimos obtener cristales derivados de selenio-metionina, y tras ser sometidos a radiación con rayos X, se obtuvo un patrón de difracción similar al expuesto en la sección de *Resultados* (Figura 42), que no pudo ser procesado. Descartamos un posible problema en la vitrificación como causante del deterioro en la difracción, con lo que pudimos concluir que era debido al desorden intrínseco del cristal originado por la sustitución del azufre de la metionina por el átomo de selenio, en esta condición de cristalización. En la actualidad se está procediendo a la obtención de cristales derivados de átomos pesados diferentes a los ya usados, como son el Gadolinio, o sales que incluyen yodo y bromo. Ha sido posible la recolección de varios juegos de datos de dichos derivados, y se está trabajando en el procesamiento y resolución de la estructura.

La información conseguida tras resolver la estructura de primer motivo “coiled-coil” representa el primer, y único hasta el momento, resultado estructural obtenido y publicado para la proteína BRMS1. Es importante mencionar que el tamaño del motivo “coiled-coil” cuya estructura se ha resuelto puede ser suficiente para desempeñar su función biológica. Algunos ejemplos de estructuras tridimensionales de motivos “coiled-coil” que encontramos en la base de datos “PDB” pertenecientes a proteínas de gran relevancia biológica, son el de un inhibidor de la replicación del ADN (Geminina) que regula el estado de oligomerización de la proteína y con ello su actividad inhibitoria sobre la síntesis del ADN (PDB = 1t6f), o un polipéptido sintético de 17 residuos que forma un “coiled-coil” que a temperaturas elevadas agrega en fibrillas amiloideas (PDB = 1s9z). Con ello, observamos la importancia que pueden adquirir fragmentos de pequeño tamaño, de tan solo 1,9 % de la secuencia total, como es el caso de la proteína APC (PDB = 1deb), de la Coronina, donde el motivo “coiled-coil” representa el 7 % de la proteína (PDB = 2akf), de GCN4, con 11,4 % (PDB = 1ij3), o de la SP-D (proteína D surfactante), con un 10,5 % del total (PDB = 1m7l). En nuestro caso, el fragmento 51 al 84 de la proteína BRMS1 ocupa aproximadamente el 14 % de la molécula total, situándose sobre la media de longitudes de motivos “coiled-coil” encontradas en la base de datos.

En este sentido, cabe señalar que hemos puesto de manifiesto la importancia del primer motivo “coiled-coil” de la proteína BRMS1 en su interacción con NMI. Las funciones biológicas que ejerce BRMS1 van a venir condicionadas por su capacidad de establecer interacciones, probablemente a través de dicho motivo, con otras proteínas, como las involucradas en metástasis, que abarcan un elevado número (DeWald *et al.*, 2005; Hurst *et al.*, 2006; Hurst *et al.*, 2008; Meehan *et al.*, 2004; Meehan y Welch, 2003;

Saunders *et al.*, 2001) y en regulación de NFκB (Liu *et al.*, 2006; Samant *et al.*, 2007; Yang *et al.*, 2008).

Con la interacción demostrada entre las proteínas BRMS1 y el co-factor de la transcripción NMI, establecemos un vínculo directo del supresor de metástasis con procesos transcripcionales, que ya venían describiéndose anteriormente en diversos estudios donde se observaba un efecto por parte de BRMS1 sobre la inhibición de la transcripción.

2. Tax1BP1 y su papel en mecanismos de ubiquitinación

La proteína Tax1BP1, de 789 residuos, ha sido asociada con diferentes tipos de enfermedades, tales como la valvulitis inflamatoria cardíaca, en ratones defectivos de la proteína (Iha *et al.*, 2008). Además, al ser una proteína de unión a la oncoproteína Tax, que es un factor de transcripción del retrovirus HTLV-1 asociado a la leucemia adulta de células T (Matsuoka, 2005), o a la enfermedad neuroinflamatoria HAM-TSP, posiblemente exista una estrecha relación de Tax1BP1 con estas enfermedades. Tax también tiene actividad represora sobre la expresión de determinados genes celulares, de gran relevancia biológica (Arnulf *et al.*, 2002; Doucas y Evans, 1999; Jeang *et al.*, 1990; Kibler y Jeang, 2001; Nicot *et al.*, 2001; Pise-Masison *et al.*, 1998; Riou *et al.*, 2000; Van *et al.*, 2001).

Se ha descrito, tanto que la proteína Tax, como Tax1BP1 están involucradas en la vía de activación del factor de transcripción NFκB, la primera activándola (Yoshida, 2001) y la segunda inhibiéndola (Iha *et al.*, 2008), por lo que la relevancia de Tax1BP1 residiría principalmente en esta ruta. Otro aspecto a tener en cuenta para analizar el papel de Tax1BP1 en este proceso es la relación que tiene con mecanismos de ubiquitinación vinculados con el estado de activación e inhibición de diferentes componentes que actúan sobre la vía. De esta forma, Tax1BP1 forma parte de un complejo multiproteico, integrado por las proteínas Itch, Tax1BP1 y A20, que parecen inhibir la ubiquitinación de TRAF6 y RIP1 por reclutamiento, y por ello regula negativamente procesos inflamatorios y de activación de la transcripción mediada por NFκB. Además, la formación de este complejo es inhibida por Tax (Shembade *et al.*, 2008).

Los procesos de ubiquitinación cumplen un papel muy importante en la regulación de esta vía de señalización y en otras rutas de señalización subcelulares. Hay estudios realizados que relacionan la ubiquitinación como un mecanismo necesario para la activación de la ruta NFκB (Chen, 2005; Haglund y Dikic, 2005) y lo que es clínicamente mas importante, como un mecanismo relacionado con cáncer (Hoeller *et al.*, 2006). Tanto moléculas de ubiquitina, como proteínas tipo ubiquitinas, actúan como mensajeros de señales que controlan muchas funciones celulares, como proliferación celular, apoptosis, ciclo celular y reparación del ADN.

Examinando las interacciones establecidas por la proteína Tax1BP1 con otras moléculas que interaccionan en procesos de ubiquitinación, así como la presencia en dicha proteína de dos motivos predichos de unión a ubiquitina estructurados en “dedos de zinc”, comprendemos que Tax1BP1 cobra una gran relevancia biológica en mecanismos de

ubiquitinación, regulando así la actividad o funciones de muchas moléculas. Además, se sabe que las alteraciones o la des-regulación de las vías de ubiquitinación resultan en el desarrollo de enfermedades humanas, incluyendo el desarrollo de tumores (Haglund y Dikic, 2005; Hoeller *et al.*, 2006).

2.1 Relevancia biológica de los dominios UBZ de Tax1BP1

Si analizamos las interacciones de Tax1BP1 con las demás moléculas involucradas en el complejo, observamos que la unión a RIP1 es a través de los dos dominios “dedos de zinc” (UBZs) y que la interacción con la ligasa de ubiquitina Itch es a través de los dominios conservados PPxY de Tax1BP1 (Shembade *et al.*, 2008), siendo del tipo WW(Itch)-PPxY(Tax1BP1). Con ello, Tax1BP1 funcionaría como adaptador para reclutar RIP1 sobre el complejo formado por la ligasa de ubiquitina y A20, con el fin de eliminar las ubiquitinas unidas a RIP1, e impedir su función activadora, por lo que no se descarta la presencia de alguna molécula de ubiquitina como intermediaria. El bloqueo sobre la regulación positiva llevada a cabo por TRAF6, de la ruta de activación de NFκB podría ocurrir de la misma manera que el expuesto para RIP1 (Figura 7).

Por ello, los dos dominios “dedos de zinc” Cys₂/His₂ de Tax1BP1 predichos como de unión a ubiquitina (UBZ o “UBiquitin Zinc finger”) (Iha *et al.*, 2008), desempeñarían un papel muy importante en los mecanismos de ubiquitinación, mediando la unión a ubiquitinas y con ello en el contexto biológico de la proteína. Este fue el motivo por el que decidimos centrarnos en el estudio estructural de ambos dominios.

Se han realizado recientemente estudios para comprobar la hipotética unión a ubiquitina de los dos motivos en “dedos de zinc” UBZ1 y UBZ2 de Tax1BP1 localizados en su extremo carboxilo-terminal. Mutaciones realizadas sobre ambos dominios muestran que la unión de la proteína completa con ubiquitina es inhibida de forma mas fuerte al mutar el segundo UBZ. Del mismo modo, muestran que el dominio UBZ2 es suficiente para la unión a cadenas de poli-ubiquitina (Iha *et al.*, 2008).

Con el fin de caracterizar estructuralmente estos dominios “dedos de zinc”, y estudiar las diferencias estructurales entre ellos que marcan la afinidad en la unión a ubiquitina, diseñamos tres construcciones: UBZ1, UBZ2 y UBZ1+2.

2.2 Estudios estructurales de UBZ1+2

Conseguimos resolver la estructura de los dos “dedos de zinc”, el fragmento comprendido entre los aminoácidos 725 a 789 (UBZ1+2). Para poder asignar las fases, en un principio centramos todos los esfuerzos en intentar conseguir un cristal derivado de alguno de los átomos pesados, oro, platino o mercurio, pero nos percatamos de la existencia de señal anómala en el cristal nativo, lo que nos hizo sospechar que algún átomo intrínseco del cristal podría estar generando dicha señal. Analizamos la posibilidad de que esta señal procediera de los átomos de zinc coordinados entre sus dos cisteínas y dos histidinas, de la forma Cys₂/His₂.

Aunque la longitud de onda usada en la recolección (0,934 Å) era lejana de la energía óptima para el zinc donde se maximizan las diferencias entre los factores de dispersión real e imaginario (f' y f'' respectivamente), obtuvimos suficiente diferencia anómala como para resolver las fases usando un sólo cristal (*SAD*), mediante el programa *SHARP*. A continuación, se expone un diagrama de la longitud óptima a la que tras excitar los electrones de los átomos de zinc, muestran mayor diferencia entre f' (-8,1) y f'' (3,9), esta es a 1,28 Å. En el mismo diagrama se muestran (con asteriscos) los parámetros que usamos en la recolección de datos y la diferencia de f' y f'' experimental, -0,07 y 2,28 respectivamente (Figura 81).

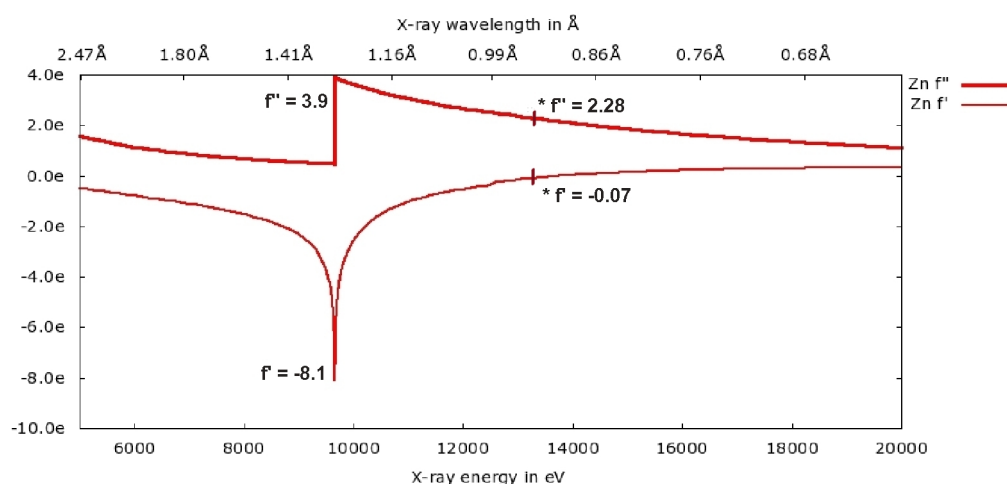


Figura 81. Diagrama mostrando las diferencias en los factores de dispersión real e imaginario (f' y f'') del átomo de zinc al usar diferentes longitudes de onda y con ello diferentes energías. La longitud de onda a la cual la diferencia entre f' y f'' es máxima para el zinc es de 1,28 Å, emitiendo a una energía de 9,66 keV. Como $*f'$ y $*f''$ se exponen los factores de dispersión de electrones, obtenidos en nuestro experimento, al usar una longitud de onda lejana a la óptima, ésta es de 0,934 Å (13,274 keV).

Finalmente, el modelo obtenido, que contenía 11 moléculas en la unidad asimétrica dispuestas en paralelo (Figura 64-A), mostraba dos átomos de zinc por cadena, coordinados por dos histidinas y dos cisteínas de cada *UBZ* (Figura 65). Los 22 átomos de zinc de la unidad asimétrica tienen una ocupación del 100 %. Los valores de los factores de temperatura de las 11 cadenas varían de 20 a 45 Å² entre los átomos de una misma cadena, correspondiendo a valores mas bajos a los que se sitúan en el centro de la cadena, y a valores mas elevados los situados en los extremos. Estos átomos terminales quedan muy expuestos al solvente, sin poder establecer contactos entre átomos próximos, por lo que tienen mayores *factores B* que los centrales, que se empaquetan al establecer contactos entre las cadenas colindantes.

Cada dominio *UBZ* consiste en una estructura en α -hélice que abarca los residuos Gln23 a His33 y Gln50 a Asn64 respectivamente, que se empaqueta contra el resto de la estructura a través de contactos con el átomo de zinc, como se observa en la Figura 82. El *UBZ2* contiene además dos pequeñas láminas- β antiparalelas, entre los residuos Lys35 a Cys37 y Ser41 a Phe44. Con idea de estudiar las diferencias conformacionales entre los dos dominios, se superpusieron ambas estructuras (Figura 82). Se observa una diferencia en la longitud de la hélice, siendo 4 aminoácidos mas larga la del *UBZ2* que la del *UBZ1*.

Aunque la superposición revela que el plegamiento de ambos dominios es similar, únicamente la estructura del *UBZ2* adoptaría una configuración clásica en “dedo de zinc” del tipo C_2H_2 , con dos pequeñas β -láminas y una α -hélice en el extremo carboxilo-terminal. Esta disposición $\beta\beta\alpha$ está ausente en el *UBZ1*, que carece de las lámina- β .

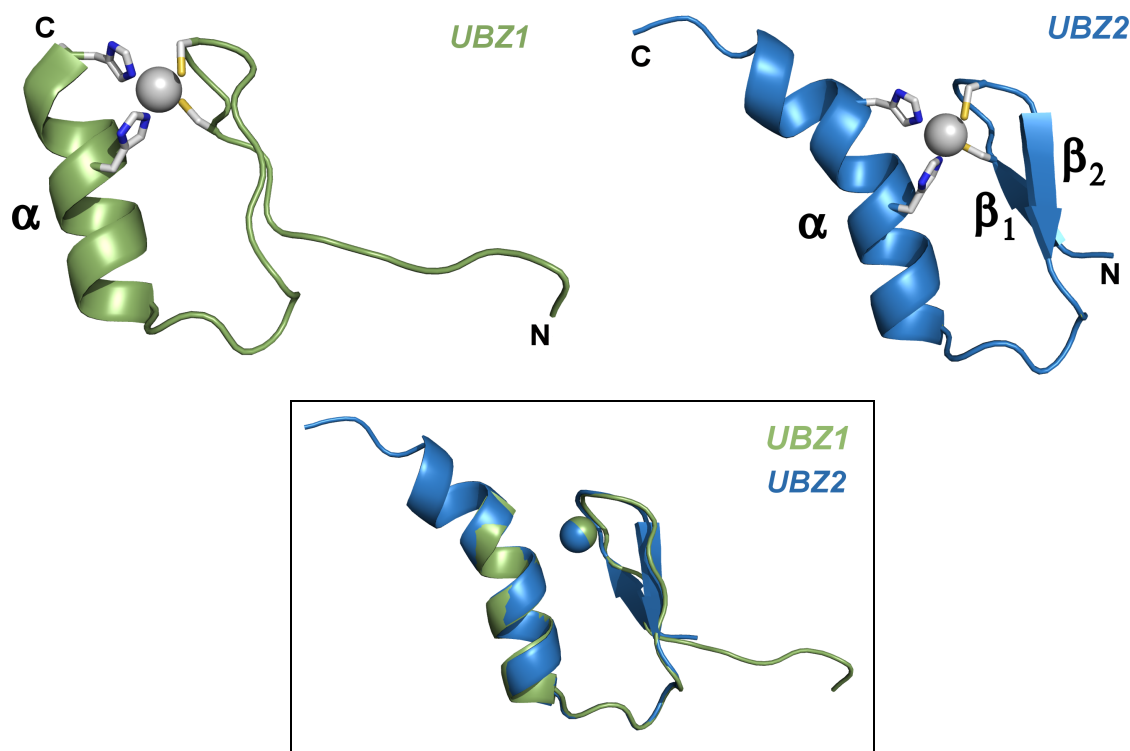


Figura 82. Estructura de los dominios “dedos de zinc” de Tax1BP1, mostrando la coordinación con los átomos de zinc del tipo C_2H_2 . En el panel de arriba se muestran en color verde y azul el *UBZ1* y *UBZ2* respectivamente, y en el panel de abajo, recuadrado, la superposición de ambos dominios. Las estructuras en α -hélice y β -lámina se representan como α y β . Los extremos amino- y carboxilo- terminales se muestran como “N” y “C” respectivamente.

Finalmente, es importante mencionar que con la resolución de la estructura tridimensional de ambos dominios *UBZ*, comprobamos la veracidad de sus límites predichos. Tras comparar la longitud de ambos dominios, observamos que el *UBZ2* definitivamente ocupa al menos 4 residuos mas en su extremo carboxilo-terminal, por lo que la predicción mostrada en la Figura 4, hasta el aminoácido 780, no coincide con la longitud real de la α -hélice tridimensional, que ocupa hasta el residuo 784 (Figura 84).

2.3 *UBZ, como dominio de unión a ubiquitina*

Estudios realizados recientemente muestran que la interacción entre Tax1BP1 y ubiquitina está mediada por el extremo carboxilo-terminal de la proteína, que contiene ambos dominios “dedos de zinc” C_2H_2 . En dichos estudios demuestran que la unión a mono-ubiquitina es inhibida de manera mas fuerte cuando se muta el *UBZ2*, y que dicho dominio también interacciona con cadenas de poli-ubiquitina (Iha *et al.*, 2008). Análisis bioinformáticos muestran que el *UBZ2* pertenece a la familia de estructura en “dedo de

zinc” de unión a ubiquitina (*UBZ*), identificada originalmente en la Polimerasa η (Bienko *et al.*, 2005).

Al menos se han identificado hasta el momento 15 tipos de dominios de unión a ubiquitina (UBD; “Ubiquitin Binding Domain”) diferentes. Por otro lado, son escasos los dominios estructurados en “dedo de zinc” que unen a ubiquitina descritos, entre ellos están los dominios *NZF* (“Npl4 Zinc Finger”), los dominios *PAZ* (“Piwi Argonaute Zwiller”) (Hicke *et al.*, 2005; Hook *et al.*, 2002; Seigneurin-Berny *et al.*, 2001), el dominio *RUZ* (“Rabex-5 Ubiquitin Binding”) (Penengo *et al.*, 2006) o los *UBZs* (Bienko *et al.*, 2005). Los tres primeros son del tipo Cys4 (C₂-C₂), mientras que el dominio *UBZ* descrito, de la familia de la polimerasa η responde a C₂H₂. Normalmente los dominios reconocen ubiquitina a través de su estructura helicoidal, pero se conocen otros tipos de reconocimiento como es el caso del dominio *NZF* o los *PAZ*, que interacciona a través de residuos localizados en los bucles expuestos (Harper y Schulman, 2006; Hurley *et al.*, 2006). Recientemente se ha descrito la estructura “dedo de zinc” del tipo C₃H de la proteína NEMO como dominio funcional de unión a ubiquitina, que interacciona a través de residuos de la hélice con la superficie hidrofóbica de la ubiquitina.

Además, existe una divergencia en la forma de reconocimiento de la ubiquitina. Se han descrito hasta el momento tres superficies de reconocimiento localizadas en dicha molécula: la hidrofóbica, que localiza a la Ile44 como residuo principal de interacción (canónica), una segunda superficie polar, centrada en el Asp58 y una última a través del extremo carboxilo-terminal, mediado por Arg74-Gly75-Gly76 (Harper y Schulman, 2006; Hurley *et al.*, 2006). Aunque la manera típica o mas consensuada de reconocimiento es la mediada por la forma canónica, llevada a cabo por *NZF*, *PAZ* y *UBZ* (Alam *et al.*, 2004; Bienko *et al.*, 2005; Bomar *et al.*, 2007; Hicke *et al.*, 2005), el dominio *RUZ* une a ubiquitina por el Asp58 (Penengo *et al.*, 2006), y otros dominios UBD como el “dedo de zinc” C₃H de UBP (“Ubiquitin-specific Protease”) lo hacen a través de las dos glicinas (R74-G75-G76) situadas en el extremo carboxilo de la ubiquitina (Bonnet *et al.*, 2008; Reyes-Turcu *et al.*, 2006).

El dominio *UBZ* de la polimerasa η es el primer “dedo de zinc” con coordinación C₂H₂ implicado en unión a ubiquitina (Bomar *et al.*, 2007). Los dominios “dedos de zinc” de Tax1BP1, representaría el segundo motivo con estas características descrito hasta el momento. Con el fin de intentar identificar los residuos que estarían implicados en la interacción canónica con ubiquitina, realizamos una comparación con diferentes patrones secuenciales presentes en otros UBDs. Para ello utilizamos un alineamiento realizado por diferentes autores, quienes han intentado establecer una secuencia consenso localizada en la α -hélice que ayudase a esclarecer la interacción *UBZ*(poli- η) con Ubiquitina, mediada únicamente por la estructura helicoidal (Bomar *et al.*, 2007). Si comparamos con la secuencia helicoidal de otros dominios que unen a ubiquitina a través de su hélice, como son UIM (“Ubiquitin-interacting motif”) (Hofmann y Falquet, 2001; Swanson *et al.*, 2003) o MIU / IUIM (“Motif interacting ubiquitin”) de Rabex-5 (Penengo *et al.*, 2006), observamos que el único residuo que se mantiene invariable entre los tres tipos de dominios es una alanina (A656 del *UBZ* de poli- η). Además, un residuo de leucina localizado tres posiciones a la derecha de la alanina (+3), está ampliamente conservada en los motivos MIU/IUIM, al igual que residuos hidrofóbicos a la izquierda de la misma y un aspártico situado en la posición -4 (Bomar *et al.*, 2007) (Figura 83-A). Ensayos de mutación de la alanina demuestran que juega un papel crucial en la interacción con la ubiquitina. Muy recientemente, se ha descrito el dominio “dedo de zinc” de NEMO como

otra estructura tipo $\beta\beta\alpha$, emparentada con el *UBZ* de poli- η . Su superficie de contacto, al igual que para el *UBZ*, sería a través de la hélice, respondiendo a una interacción canónica, con cinco residuos (posiciones -7, -4, -3, -1 y +3) conservados, mientras que la alanina central se reemplaza por una metionina (Figura 83-B) (Cordier *et al.*, 2008).

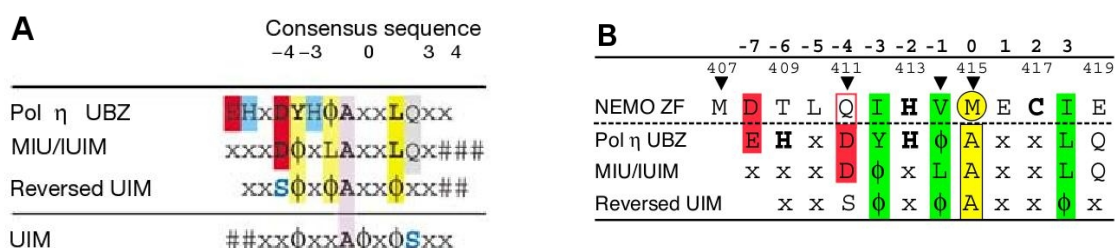


Figura 83. Alineamiento de las secuencias consenso entre diferentes tipos de dominios de unión a ubiquitina. **A)** Alineamiento del dominio *UBZ* de la polimerasa η con MIU/IUIM y UIM reverso. La alanina central invariante (posición 0) se representa en color púrpura, un residuo de glutamina altamente consensuado en gris, los residuos hidrofóbicos conservados en amarillo, los ácidos en rojo y los que se coordinan con el átomo de zinc en azul, así como una serina en la posición +4 en UIM, que es reemplazada por un aspártico en los dominios MIU/IUIM y *UBZ* (Bomar *et al.*, 2007). **B)** Alineamiento de la secuencia del “dedo de zinc” de NEMO con las secuencias consenso del *UBZ* de poli- η , MIU/IUIM y UIM reverso. El aminoácido altamente conservado alanina se muestra en amarillo, y es reemplazado por una metionina en NEMO. Residuos conservados hidrofóbicos son representados en verde, y los ácidos en rojo. La coordinación con los átomos de zinc se expone en negrita. Los triángulos aluden a las mutaciones analizadas en este estudio (Cordier *et al.*, 2008). ϕ y # representan residuos hidrofóbicos y ácidos respectivamente, y “x” significa cualquier aminoácido.

Si analizamos la existencia en los “dedos de zinc” de Tax1BP1 de este consenso establecido en los anteriores estudios, observamos que no existe ninguna similitud en la secuencia de la α -hélice. En la Figura 84, donde se muestra la secuencia de ambos dominios, observamos que en la posición 0, no están presentes ni el residuo de alanina ni la metionina. Únicamente los residuos posicionados en -7 (Glu) y -4 (Gln) en los *UBZ1* y *UBZ2*, se presentan conservados con respecto a la poli- η y NEMO respectivamente, que no serían suficientes para comparar estructuralmente la interacción, pues no forman parte del contacto con la ubiquitina.

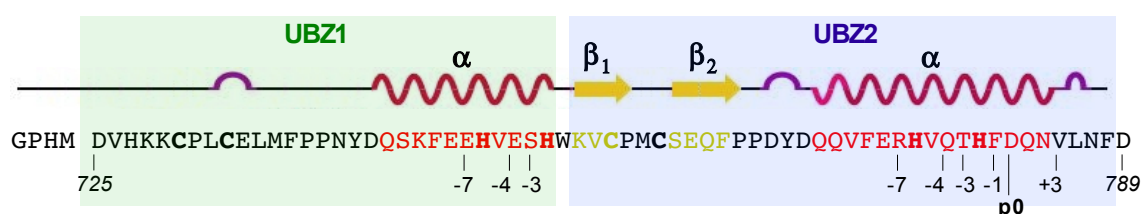


Figura 84. Representación de los dominios de unión a ubiquitina de la proteína Tax1BP1, mostrando en un diagrama su estructura secundaria obtenida tras la resolución de la estructura. Los dominios *UBZ1* y *UBZ2* se muestran sobre fondo verde y azul respectivamente. La secuencia correspondiente a la estructura en α -hélice (α) se muestra en color rojo, mientras que las lámina β (β_1 y β_2) se representan en amarillo. La numeración inicial y final expuesta, corresponde a los aminoácidos en la secuencia de Tax1BP1. Se pueden también observar la asignación en las posiciones consenso establecidas mediante estudios previos (Bomar *et al.*, 2007), siendo “p0” la inicial, ocupada por alanina en la secuencia consenso de los dominios de unión a

ubiquitina de poli-η, IUIM y IUIM reverso y las posiciones -1, -3, -4, -7 y +3 conservadas entre dichos dominios de las proteínas poli-η y NEMO.

La posición 0 está ausente en el *UBZ1*, que dispone de una hélice mas corta que el segundo dominio en su extremo carboxilo, por lo que la región donde se localiza la secuencia consenso ni siquiera está presente. Probablemente ésta sea la causa por la que la interacción del dominio *UBZ1* con ubiquitina no sea tan afín como la que presenta el dominio *UBZ2*, que tridimensionalmente presenta una elevada homología con el *UBZ* de poli-η, tal y como se muestra en la Figura 85, a pesar de no compartir su secuencia consenso de unión. El dominio *UBZ* de la polimerasa η contiene, al igual que en la proteína Tax1BP1, dos pequeñas β-láminas que contactan con la α-hélice a través del átomo de zinc. La interacción con la molécula de ubiquitina en este dominio descrito sería a través de la alanina 33, mostrada en la figura, que se expone hacia la cara externa de la hélice.

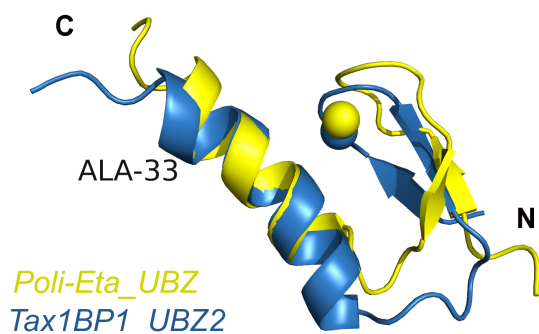


Figura 85. Superposición de las estructuras tridimensionales de los dominios “dedos de zinc” del tipo C₂H₂ de unión a ubiquitina *UBZ2* de Tax1BP1 (azul) y de la proteína polimerasa η (PDB: 2150) (amarillo). Los átomos de zinc se representan como esferas del mismo color que la molécula a la que corresponden. Las estructuras α-hélice se muestran con conformación helicoidal, mientras que las β-láminas se exponen con flechas. La semejanza mostrada en la superposición entre sus extremos amino- (N) y carboxilo- (C) terminales muestra una elevada homología estructural.

Se identifica el residuo Ala33, que parece desempeñar un papel crucial en la interacción de la polimerasa η con ubiquitina.

Esta diferencia en la conservación del residuo de alanina referido, hace pensar que si la interacción con ubiquitina tiene lugar a través de la α-hélice, se estaría llevando a cabo de forma distinta al del *UBZ* de poli-η, además de estar mediada posiblemente por otros residuos. Este modo de interacción diferente, dotaría a la estructura del dominio *UBZ2* de la proteína Tax1BP1, de una relevancia aún mayor por ser el primer “dedo de zinc” del tipo C₂H₂, que reconoce ubiquitina mediante un mecanismo aún no descrito.

Por otro lado, cabe la posibilidad de que el reconocimiento con la ubiquitina se lleve a cabo por otra zona diferente a la α-hélice. Como se mencionó anteriormente, está descrito que otros dominios de unión a ubiquitina pueden interaccionar con ella a través de residuos localizados en otras zonas, como es el caso del dominio *RUZ* de la proteína Rabex-5, que a través de los residuos Y25-Y26-N28-C35-S36 reconoce la superficie polar centrada en Asp58 de la ubiquitina (Penengo *et al.*, 2006), o el dominio *NZF* de Npl4 que de manera canónica reconoce a la ubiquitina por los residuos T13-F14-M25 (Alam *et al.*, 2004), ambos estructurados en “dedos de zinc” del tipo C₄. Para estudiar esta posibilidad, se realizó un alineamiento entre el dominio *UBZ2* de Tax1BP1 y el dominio *NZF*, que es el que mostraba mas homología tridimensional con los dominios *UBZs* (Figura 86), con el fin de comparar las estructuras y la conservación de los residuos de interacción con ubiquitina. En la unión mediada por *NZF*, los residuos T13 y F14 parecen desempeñar el

papel mas importante en la interacción, aminoácidos que están reemplazados por S41 y E42 en el *UBZ2* de Tax1BP1. Además, sus cadenas laterales se localizan en las mismas posiciones que lo hacen en el dominio *NZF*, como se observa en una superposición de ambas estructuras mostrada en la Figura 86.

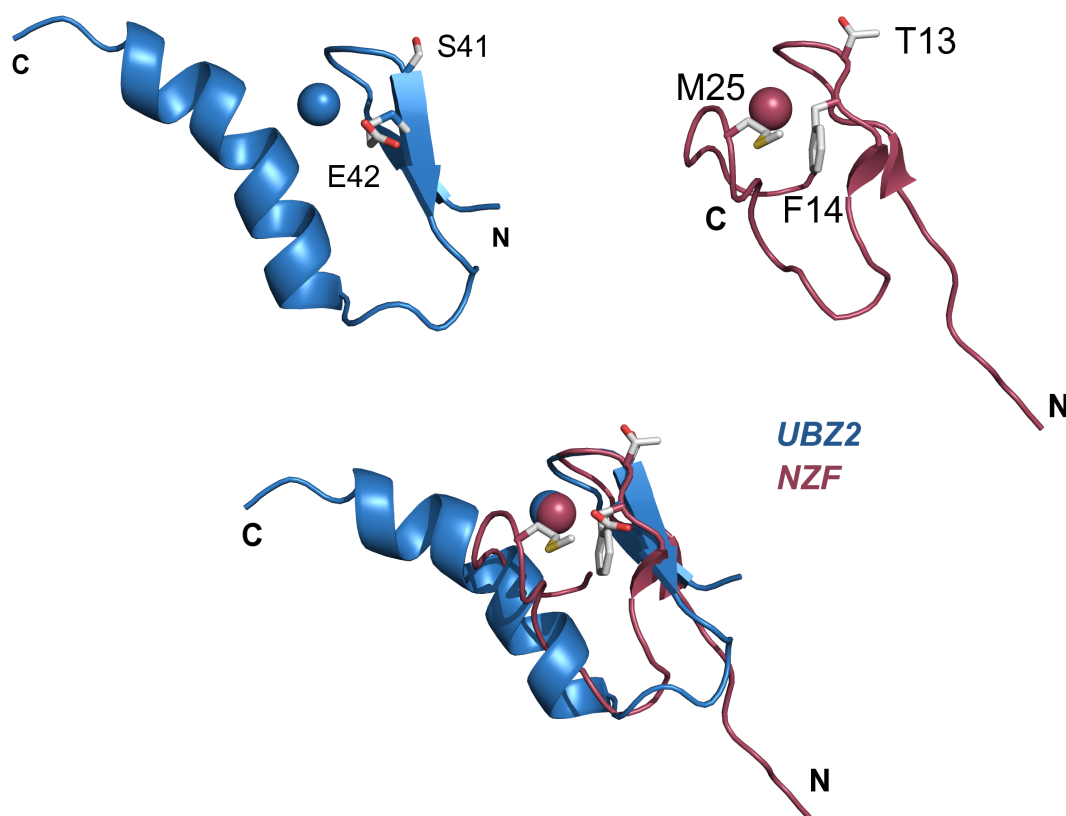


Figura 86. Superposición de los dominios “dedos de zinc” del tipo C_2H_2 y C_4 correspondientes a las proteínas Tax1BP1 (*UBZ2*; en color azul) y Npl4 (*NZF* (PDB: 1Q5W); en color rojo). La α -hélice se representa en conformación helicoidal, y las β -lámina como flechas. Los residuos del dominio *NZF* cuyas cadenas laterales se detallan (T13, F14 y M25) corresponden a los que está descrito intervienen en la interacción con ubiquitina, y los representados en el dominio *UBZ* (S41 y E42), los que sustituirían a los anteriores en el caso de responder al mismo tipo de unión a ubiquitina en este modelo. “N” y “C” corresponden a los extremos amino- y carboxilo- terminales.

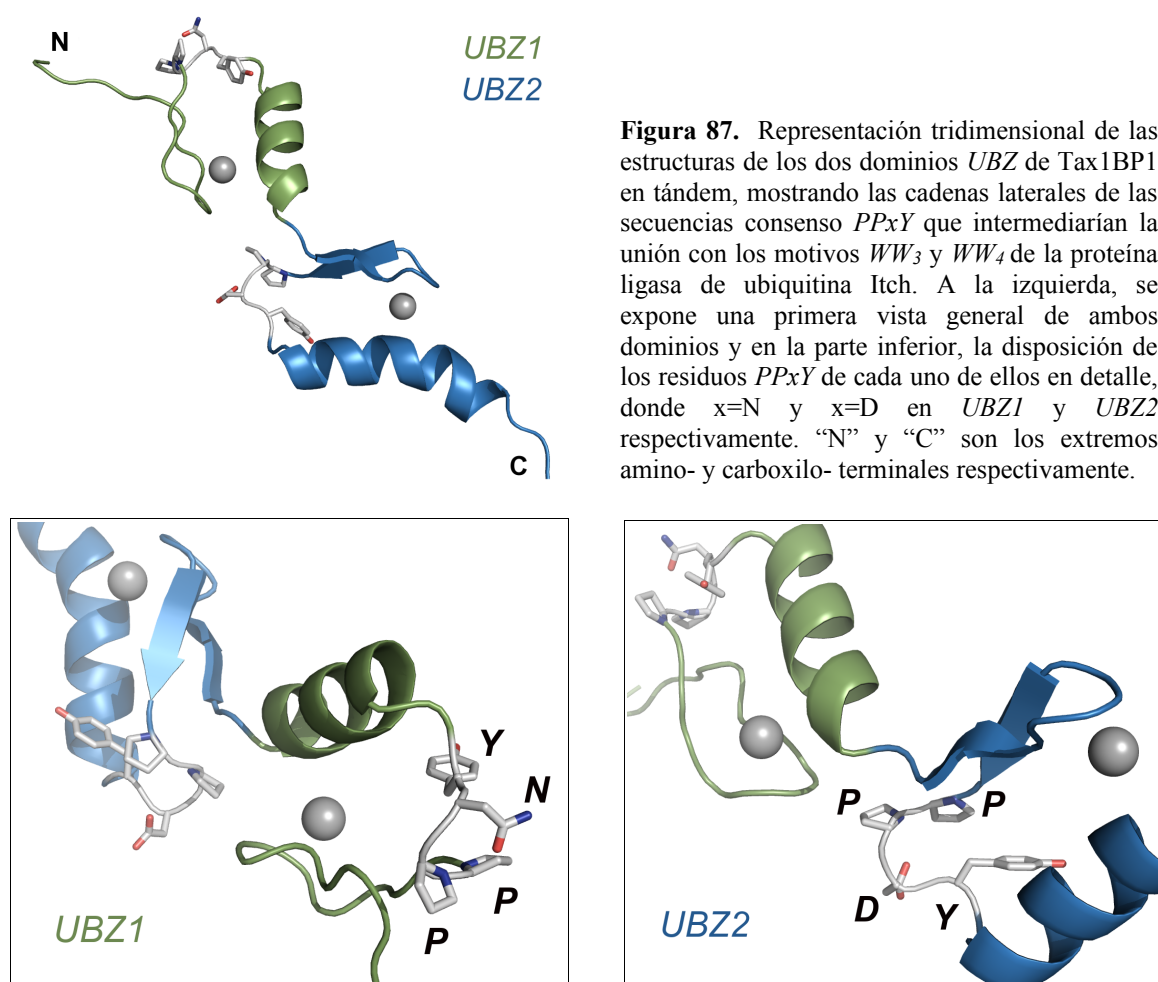
Aunque, pese a obtener cristales, no conseguimos resolver la estructura de los dominios *UBZ1+2* en presencia de ubiquitina y de poli-ubiquitina, sí obtuvimos indicios de la formación de los dos complejos bajo las condiciones usadas en la cromatografía de exclusión por tamaño molecular (Figuras 68 y 69). Si la interacción entre el dominio *UBZ2* y la proteína ubiquitina fuera caracterizada estructuralmente, podríamos verificar si se trata del primer caso de reconocimiento de un dominio *UBZ* del tipo C_2H_2 a través de residuos no localizados en su α -hélice.

Los resultados obtenidos en estos estudios estructurales representan los primeros descritos en relación a la proteína Tax1BP1, además de ser muy novedosos en cuanto a la resolución de la estructura de un “dedo de zinc” *UBZ* que interacciona con ubiquitina, ya que existe únicamente un modelo de interacción de dominio *UBZ* (el de la polimerasa η) con ubiquitina (Bienko *et al.*, 2005).

2.4 Interacción de UBZ1+2 con la ligasa de ubiquitina Itch, a través de las secuencias consenso PPxY

La ligasa E3 de ubiquitina Itch es una proteína que forma parte del complejo multiproteico dirigido por la molécula A20, implicado en procesos de ubiquitinación en la ruta de activación de NFκB (Shembade *et al.*, 2008). Mediante cromatografía de exclusión por tamaño molecular (Figura 72) obtuvimos también evidencias de la interacción entre las secuencias consenso PPxY, situadas en cada dominio UBZ de la proteína Tax1BP1, y los dominios ricos en triptófanos de Itch (WW). Está predicho que secuencias consenso PPxY interaccionan con dominios WW (Shembade *et al.*, 2008; Sudol *et al.*, 1995). Concretamente, la proteína estudiada correspondía a los dominios WW₃ y WW₄, que según nuestros resultados, serían suficientes para mediar la interacción con ambos dominios UBZ de Tax1BP1.

Al analizar la disposición tridimensional de esta secuencia consenso PPxY, situada antes de la α-hélice en ambos dominios, y empaquetada entre la lámina β₂ y la hélice en UBZ2, observamos que se localiza en un lazo expuesto hacia la cara exterior de la estructura (Figura 87), donde la superficie es de carácter hidrofóbica principalmente.



Gracias a los resultados obtenidos en este trabajo, podemos especular sobre el modo de interacción entre Itch y Tax1BP1. La longitud de los dos dominios *WWs* de Itch, equiparable a los *UBZs* de Tax1BP1, hace pensar que la interacción se esté dando de manera UBZ1-WW, UBZ2-WW, interaccionando cada dominio *UBZ* con cada uno de los dominios *WW*.

La descripción de la interacción con una proteína ligasa de ubiquitina, fortalece mas la idea de la formación de un complejo de regulación de procesos de ubiquitinación que controla rutas importantes, como la ruta de NFκB. A ello se suma los resultados obtenidos recientemente que demuestran la regulación negativa que Itch ejerce sobre dicha ruta (Shembade *et al.*, 2008).

Conclusiones

- La resolución, mediante Cristalografía de Rayos X, de la región comprendida entre los residuos 51 a 84 de la proteína humana supresora de metástasis BRMS1 demuestra que se trata de un motivo “coiled-coil” levógiro antiparalelo, como así lo corrobora la asignación del patrón de “heptada” típico de estos motivos.
- Los resultados obtenidos en los experimentos de resonancia magnética nuclear y de calorimetría de titulación isotérmica, confirman la interacción entre el primer motivo “coiled-coil” de la proteína BRMS1 y el motivo predicho como “coiled-coil” de NMI, con una afinidad del orden de μM .
- Se ha resuelto la estructura de los dos dominios “dedos de zinc” en tándem de la proteína humana Tax1BP1, predichos como dominios de unión a ubiquitina. La comparación estructural de ambos con otros dominios de las mismas características, evidencia que únicamente *UBZ2* se ajusta al plegamiento típico de los dominios “dedos de zinc” C_2H_2 , aunque no a la secuencia consenso establecida en los dominios *UBZ*, por lo que la interacción a ubiquitina estaría mediada por una región diferente. La estructura del dominio *UBZ1* muestra que no está presente la región descrita necesaria para unir a ubiquitina en este tipo de “dedo de zinc”, lo que conlleva que, de unir a esta molécula, lo haría de manera distinta, representando un nuevo motivo de unión aún no descrito.
- Obtuvimos evidencias que sugieren que el fragmento que engloba los dos dominios *UBZ* de Tax1BP1 interaccionan con ubiquitina y poli-ubiquitina. De la misma manera, también conseguimos evidencias de la interacción entre los dos *UBZs* en tándem con dos motivos ricos en triptófanos (WW_3+WW_4) de la ligasa de ubiquitina Itch.

Anexo

1. Resolución de estructuras mediante Cristalografía de Rayos X

Un requisito indispensable para hacer posible la resolución de estructura tridimensional mediante Cristalografía de macromoléculas es, como su nombre indica, conseguir un cristal de dichas macromoléculas. Con ello se obtiene una repetición ordenada en tres dimensiones de la celda unidad de moléculas, que forma la trama de la estructura cristalina. Estos cristales han de ser únicos y de buena calidad para facilitar el procesamiento y la interpretación de los datos de ellos obtenidos.

La técnica empleada en el presente trabajo para realizar estos ensayos de cristalización fue la “difusión de vapor” entre una gota, formada por la mezcla, en determinada relación volumen/volumen de la proteína en solución y del agente precipitante que se encuentra en el pocillo. De este modo, la diferencia en contenido de agua entre el reservorio y la gota, y por tanto de su presión de vapor, hace que de forma gradual, el sistema tienda al equilibrio, favoreciendo así la precipitación de la proteína. En este punto, tiene lugar la formación de cristales (Ref. bibliográfica 2*). Se trata de un sistema cerrado y bajo condiciones controladas. En la formación de cristales entran en juego dos procesos: nucleación, entendiéndose como el comienzo de un cambio de estado en una región pequeña pero estable, y crecimiento del cristal a partir de estos centros de nucleación que implica la formación de una interfaz en los límites del nuevo estado (sólido y líquido, entre el cristal y el solvente).

Encontrar las condiciones exactas de cristalización para una proteína específica se convierte en un reto experimental que hace de la cristalografía un arte, más que una ciencia.

Otro gran reto tras la obtención del cristal es encontrar las condiciones óptimas para un enfriamiento rápido del mismo (a 130 K) sin que sufra daños. Al incidir sobre la muestra, la radiación de Rayos X puede producir daño directo o primario sobre cualquiera de los componentes del cristal y también daño secundario cuando alguno de los radicales generados en el primario migran hacia otras zonas de cristal y provocan a su vez una nueva lesión. Una alternativa para disminuir el daño de radiación secundario es enfriar rápidamente el cristal, a través de un proceso de vitrificación, por el cual la transformación de líquido a sólido es instantánea, evitando la formación de cristales de hielo que generen señales forma de aros sobre el patrón de difracción del cristal.

El mantenimiento del cristal a baja temperatura reduce la movilidad de los radicales formados en la radiación primaria disminuyendo el daño, sin embargo, también este proceso de vitrificación puede afectar al patrón de difracción. Por ello, es preciso encontrar la técnica óptima en cada caso para hacer el cristal lo más resistente posible a la exposición de Rayos X y que en este proceso no sufra modificaciones que perjudiquen los datos (Kriminski *et al.*, 2003). El uso de soluciones crioprotectoras, tales como glicerol, etilenglicol o polietilenglicol es necesario en algunos casos para evitar esta formación de hielo cristalino. La adición de agentes crioprotectores incrementa el proceso de congelamiento del agua en 1-2 segundos, evitando la formación de hielo si la vitrificación se realiza dentro de ese intervalo de tiempo (Ref. bibliográfica 3*, pag. 1-18).

Una vez conseguido un cristal viable para su exposición a difracción de Rayos X se va a someter a una serie de etapas, resultando en un experimento de ensayo-error, ya que tras la obtención del cristal, primer inconveniente, la resolución de la estructura puede

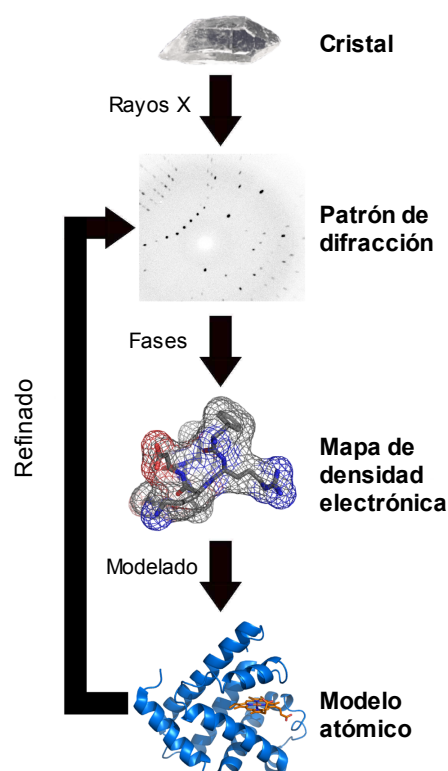
verse troncada por otra serie de dificultades. La presencia de problemas intrínsecos en el cristal, de una mala manipulación del mismo, una deficiente recolección de los datos, o equivocado procesamiento de los mismos, son ejemplos de los muchos obstáculos experimentales que son necesarios superar hasta llegar a la resolución de la estructura.

Estas etapas hasta llegar al conocimiento tridimensional de la proteína (Figura 88) se resumen en una serie de complejos pasos en los que es primordial el uso de programas cristalográficos para acceder al siguiente, etapas que serán explicadas mas adelante.

Figura 88. Representación esquemática de las etapas llevadas a cabo para la determinación de una estructura tridimensional por la técnica de cristalografía de rayos X.

A partir de la obtención de un cristal, éste es sometido a irradiación de rayos X, originando un patrón que consiste en miles de puntos, cada uno de los cuales con una posición e intensidad determinadas. Tras un procesamiento de esta difracción con programas cristalográficos se obtienen las fases de las ondas que forman cada punto y con ello un mapa de densidad electrónica sobre el que se construye un modelo estructural. Tras un proceso de refinado frente a los datos experimentales y la posterior validación del proceso, se obtiene un modelo de la estructura terciaria de la molécula en cuestión.

Esta figura ha sido adaptada de la base de datos “Wikimedia Commons”, realizada por Thomas Splettstoesser y basada en el PDB con el código 1MBO.



Los programas informáticos usados en la cristalografía para llegar al conocimiento de una estructura terciaria, engloban un amplio abanico que ofrece la posibilidad de usar diferentes métodos de resolución de la misma. El uso de un método u otro depende de los datos obtenidos y de la información existente para cada molécula, así como de las características del cristal. A continuación sólo se aludirán los métodos seguidos para la resolución de las estructuras resueltas durante la tesis.

1.1 Difracción de Rayos X

Tras la vitrificación, los cristales se irradian por un haz de alta intensidad de rayos X, que consisten en una radiación electromagnética cuya longitud de onda se encuentra comprendida entre 10^{-7} y 10^{-11} m (100 y 0,1 Å).

Al incidir sobre el cristal, el haz de rayos X se escinde en varias direcciones debido a la simetría de la agrupación de átomos, esto es: difracta, emitiendo radiación de la

misma longitud de onda en determinadas direcciones, dando lugar a un patrón de intensidades que son recogidas en el detector. Este patrón puede interpretarse a partir de la ubicación de los átomos en el cristal, aplicando la “ley de Bragg” (Ecuación 2) (Ref. bibliográfica 2*)

$$n\lambda = 2d \sin\theta$$

Ecuación 2. Ley de Bragg, donde n es un número entero, λ es la longitud de onda de los rayos X, d es la distancia entre los planos de la red cristalina y, de una familia de planos de índice hkl y θ es el ángulo entre los rayos incidentes y los planos de dispersión, o ángulo incidente (igual al de difracción), llamado “ángulo de Bragg”.

Sin embargo, el problema consiste, precisamente, en conocer la posición de los átomos en el cristal.

Estas radiaciones emitidas desde el cristal, se asemejan a haces reflectados a partir de planos paralelos y equivalentes de átomos en el cristal, por lo que a cada punto en el patrón de difracción se le llama reflexión. Cada reflexión se caracteriza por un trío de números, hkl , llamados “índices de Miller”, que identifica un grupo particular de planos equivalentes y paralelos. En una muestra debidamente ordenada se producen interferencias destructivas cuando las ondas se anulan mutuamente, sin embargo, en puntos particulares, estas ondas interfieren aditivamente cumpliendo la “ley de Bragg”, y pudiéndose predecir los ángulos en los que los rayos X son difractados por una estructura atómica periódica, podemos estudiar las direcciones en las que la difracción de rayos X produce interferencias constructivas. Una interferencia es constructiva cuando la diferencia de fase entre la radiación emitida por distintos átomos es proporcional a 2π , definido según la “ley de Bragg” (Figura 89).

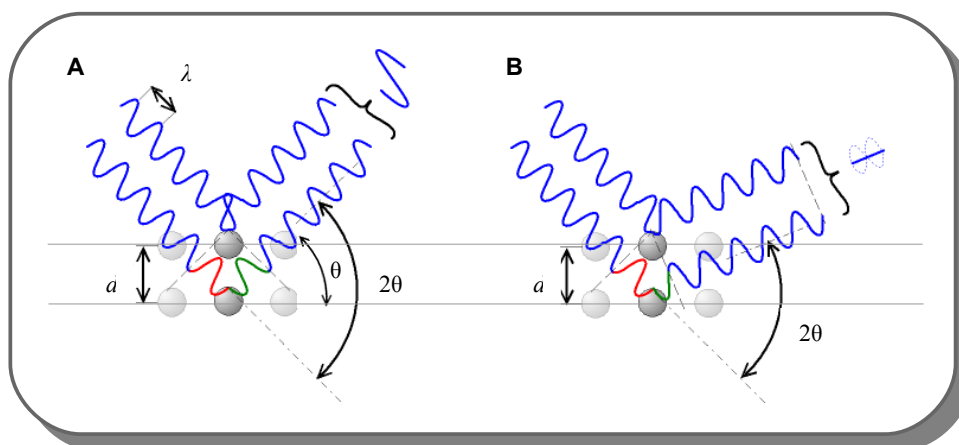


Figura 89. Esquema representativo de la deducción de la “Ley de Bragg”. Cuando el haz de rayos X llega a un átomo, este se dispersa, emitiéndose una radiación de la misma longitud de onda en determinadas direcciones. Cuando el ángulo de desviación es 2θ , el cambio de fase de las ondas produce interferencia constructiva (A), sin embargo, cuando estas ondas se anulan mutuamente se producen interferencias destructivas (B), por lo que no cumplen la “Ley de Bragg”. Se muestran todos los parámetros que conforman dicha ley. Figura modificada a partir de la base de datos “Wikimedia Commons”.

Así, un grupo de planos paralelos con un índice hkl y un espacio interplanar d_{hkl} produce un haz difractado cuando los rayos X a una longitud de onda λ inciden sobre los planos a un ángulo θ y son reflejados con el mismo ángulo, sólo si θ cumple la condición de Bragg.

1.2 Reducción de datos

Las imágenes de difracción contienen información “bruta” acerca de la organización atómica dentro del cristal. Esta información consiste simplemente en una lista de intensidades con un índice hkl asignado a cada una de ellas, que se corresponde, como se ha indicado, con el juego de planos del cristal que han generado esta reflexión. Cada reflexión puede haber sido recogida en varias imágenes consecutivas y, además, algunas reflexiones son equivalentes a otras en virtud de la simetría del cristal. Por este motivo es esencial procesar estos datos y simplificarlos para obtener la intensidad integrada de cada reflexión única, así como el error experimental en la medida (Ref. bibliográfica 3*, pag. 85-94).

En general, para la reducción de los datos, es preciso discriminar tres etapas:

- *Indexado*; donde se determinan los parámetros de la celda unidad del cristal, esto es, la longitud de sus ejes, los ángulos que la forman que definen el tamaño y la forma de la celda, así como la “red de Bravais”, entendida como una disposición infinita de puntos discretos cuya estructura es invariante bajo traslaciones, y en la mayoría de casos también bajo rotaciones o simetría rotacional. Estas propiedades hacen que desde todos los nodos de una “red de Bravais” se tenga la misma perspectiva de la red. Se dice entonces que estos puntos son equivalentes.
- *Integrado* de las intensidades de las reflexiones. Para medir la intensidad total de cada reflexión, el programa descarta reflexiones cuya señal no se diferencia del fondo, aquellas cuya posición no se ajusta a la predicción, al igual que aquellas reflexiones cuyo fondo o cuya intensidad (por ejemplo por haber sobrepasado el umbral de linealidad del detector) no pueden ser estimadas de manera fiable. Además, muy importante, el programa estima el error en cada intensidad integrada.
- *Escalado y combinado*; donde se llevan todas las reflexiones de todas las imágenes a una misma escala numérica relativa, lo que permite compararlas y combinarlas, produciendo una lista de reflexiones únicas con sus intensidades y sus errores asociados.

En el presente estudio, la reducción de datos se realizó usando el programa informático *Mosflm* para el indexado e integrado, y *Scala* del paquete informático *CCP4* (CCP4, 1994), para el escalado y combinado.

1.3 Resolución de fases experimentales

De cada una de las reflexiones, necesitamos tres parámetros para obtener el mapa de densidad electrónica. Dichos parámetros son la posición (bajo la forma del índice de Miller hkl), la amplitud (proporcional a la raíz cuadrada de la intensidad de cada reflexión, I_{hkl}) y la fase de las ondas que forman cada señal. La posición e intensidad de cada reflexión son determinadas de manera relativamente fácil, tal y como se ha expuesto anteriormente, pero las fases de las ondas que forman cada punto deben ser también determinadas, convirtiéndose en un nuevo obstáculo.

La suma de las contribuciones de todos los átomos en la celda unidad computadas para una reflexión hkl es llamado factor de estructura F_{hkl} , y representa la onda creada por superposición de muchas ondas individuales, originadas cada una de ellas por la difracción de átomos individuales. Este factor de estructura posee amplitud, frecuencia y fase. Las amplitudes de los factores de estructura (F_{hkl}) se relacionan con la densidad electrónica a través de la “transformada de Fourier”, representada por la Ecuación 3.

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l |F(hkl)| \cos 2\pi(hx + ky + lz - \phi(hkl))$$

Ecuación 3. Aplicación que responde a la transformada de Fourier. Esta función define la densidad electrónica en un punto de coordenadas (x, y, z) en la celdilla elemental. $F(hkl)$ representa los llamados factores de estructura (cuyos módulos están directamente relacionados con las intensidades de las reflexiones del espectro). “ h, k, l ” son los “índices de Miller” de las reflexiones y $\phi(hkl)$ representa las denominadas “fases” de las reflexiones (unas respecto de otras).

Esta ecuación representa la “transformada de Fourier” entre el espacio real (en donde están los átomos) representado por la función de densidad ρ y el espacio recíproco (en donde está el espectro) representado por los factores de estructura y sus fases. Las variables entre estos dos espacios son por lo tanto recíprocas (Ref. Bibliográfica 2*). La **Ecuación 3** demuestra que para obtener el valor de la densidad electrónica en un sólo punto de coordenadas (xyz) es necesario sumar las contribuciones de todos los factores de estructura.

Así, la densidad electrónica puede ser descrita como una serie de Fourier a través de los factores de estructura, que llevan información acerca de su frecuencia (determinada experimentalmente según la fuente de rayos X), su amplitud (proporcional a la raíz cuadrada de la intensidad de cada reflexión, I_{hkl}) y fase, que hay que determinar experimentalmente o bien mediante procedimientos de cálculo, siempre que haya un modelo adecuado de la molécula a resolver.

Una propiedad importante del modelo que pretendemos generar es su *resolución*. La unidad de este parámetro es el anstromg (Å), y cuanto menor sea su valor, mayor será la resolución, es decir, el mapa de densidad electrónica será de mayor calidad.

Cada átomo en la celda unidad contribuye a cada reflexión en el patrón de difracción. Así, la incorporación en sistema cristalino de un átomo específico cuya difracción aporte una pequeña perturbación en el patrón de difracción, puede ser usado

para una estimación inicial de las fases. Para que la diferencia generada por la difracción de este átomo sea percibida y medible, debe ser una difracción fuerte, por lo que hablamos de átomos pesados, esto es, con gran número de electrones, que aporten diferencias de señal. Estas diferencias, para ser explotables, pueden ser de dos tipos: bien sobre la intensidad recogida de cada reflexión, en cuyo caso se requiere que el nuevo cristal sea isomorfo, es decir con la misma celda y simetría, con el cristal nativo sin átomos pesados, o bien pueden incluir señal “anómala”, en la que la absorción temporal de la radiación por los átomos pesados induce un cambio de fase que se traduce en un cambio de la intensidad medible (Ref. bibliográfica 2*).

Los átomos “ligeros” de carbono, nitrógeno y oxígeno no contribuyen a la señal anómala en las longitudes de onda útiles para la cristalografía, pero la radiación absorbida por el átomo pesado es re-emitida con fase alterada. Así, la señal anómala altera la contribución del átomo pesado a su factor de estructura.

Una aproximación para generar señal anómala medible es obtener un cristal derivado de átomo pesado. Existen diversas técnicas para asignar las fases, de forma que cuando utilizamos la señal anómala generada en un cristal, y a una única longitud de onda, hablamos de SAD, y cuando medimos la generada a diferentes longitudes de onda, aludimos a la técnica MAD. Si a ello añadimos la información extraída de las diferencias isomorfas, nos referimos a SIR (para un juego de datos) o MIR (medida en juegos de datos de varios derivados pesados). Las combinaciones de ambas opciones originan, del mismo modo, dos nuevas técnicas: SIRAS (un derivado) y MIRAS (varios átomos pesados).

Para la resolución de las fases de las estructuras resueltas en el presente estudio, utilizamos dos de estas aproximaciones: *SIRAS* y *SAD*:

a) *SIRAS* (“*Single Isomorphous Replacement with Anomalous Scattering*”)

Para resolver las fases y con ello trazar la densidad electrónica, se usaron las diferencias anómala e isomorfa (respecto a las diferencias en amplitudes de los factores de estructura de átomos derivados y del resto del cristal) originadas por un único derivado de átomo pesado medido a la longitud de onda específica donde absorbe rayos X (para maximizar su señal anómala). Esta diferencia isomorfa es obtenida al comparar esta señal con el juego de datos recogido del cristal nativo. Se requiere que los iones de metal sean incorporados en un cristal sin perturbar la estructura y enrejado del mismo, ya que ambos juegos de datos han de ser comparados. Por ello, es crucial que conserven las dimensiones de la celda unidad y su simetría, a lo que alude el término de cristales isomorfos, lo cual en ocasiones no es posible y resolver la estructura por este método se convierte en imposible.

b) *SAD* (“*Single-wavelength Anomalous Diffraction*”)

Para resolver las fases, en este caso se utilizó exclusivamente la señal anómala recogida en un juego de datos de un cristal que mostraba la presencia de átomo pesado. En este caso no es crucial obtener un cristal nativo, ya que la señal anómala del cristal anómalo contiene toda la información necesaria para la resolución de la estructura.

1.4 Construcción y afinado del modelo

Una vez obtenido el mapa de densidad electrónica se puede empezar a trazar el modelo. Llegado a este punto, disponemos de las fases, las posiciones de los átomos pesados y la secuencia de la molécula, que es la información necesaria para llevar a cabo el modelado inicial de la estructura, mediante diversos programas cristalográficos.

Tras obtener un modelo inicial, el siguiente paso es optimizarlo para ajustarlo a los datos experimentales. Para dar una definición de afinamiento, podemos aludir a la optimización de una función de un grupo de observaciones mediante cambios en los parámetros de un modelo. Al mismo tiempo, hay que hacer consistente el modelo experimental con lo que sabemos de la química de las proteínas, de forma que se aplican restricciones estructurales sobre las distancias de los enlaces, los ángulos entre los átomos enlazados, etc, para acercar el modelo lo máximo posible a la realidad.

Con este afinado se pretende minimizar el error del modelo experimental, haciendo mínima la desviación entre lo calculado y lo teórico. En este punto, el modelo (*.pdb*) es una colección de átomos, cada uno de ellos definido por su localización (coordenadas en tres dimensiones), un índice de ocupación, que indica en qué proporción dentro del cristal ese átomo ocupa esa posición, y un factor de temperatura o de desplazamiento atómico *B*, que es una estimación de la vibración del átomo con respecto a su posición central. El *factor B* representa la disminución en la intensidad en la difracción debido al desorden del cristal, tanto estático, como el causado por vibraciones térmicas. Existen dos tipos de afinamientos: isotrópicos y anisotrópicos. En el primero se define al *factor B* con un sólo número, ya que el átomo vibra de igual manera en todas las direcciones (puede ser visualizado como una esfera), mientras que en el refinamiento anisotrópico el átomo vibra describiendo un elipsoide, por lo que el *factor B* en este caso se define con seis parámetros (Ref. bibliográfica 3*). Como la resolución de los cristales no alcanzó al punto de poder utilizar un modelo anisotrópico para cada uno de los átomos de la estructura, ya que eso hubiera conllevado tener más parámetros para refinar que observaciones medidas, se empleó la aproximación TLS (del inglés “translation-libration-screw”, ver más adelante).

Para llevar a a cabo el afinamiento de la molécula se establecieron diferentes métodos (facilitados en el programa *REFMAC5* (Murshudov *et al.*, 1997)) con distintos parámetros según la estructura:

- *Afinamiento de cuerpo rígido* (en inglés “*Rigid Body*”). El afinamiento se realiza en base a 6 parámetros, 3 de rotación y 3 de translación, viendo la estructura como un bloque o una estructura rígida que se mueve respecto a estos parámetros, es decir, sin refinar por separado los parámetros de orientación y localización de cada átomo en el cristal.
- Aplicación de *simetría no cristalográfica* (*NCS*). Se trata de una aproximación similar a la del afinado de cuerpo rígido, pero relajando las restricciones, de forma que las moléculas que, dentro de la unidad asimétrica están relacionadas por una rotación y translación de cuerpo rígido, tengan una disposición parecida aunque no necesariamente idéntica. Evidentemente, sólo se puede aplicar simetría no cristalográfica cuando la unidad asimétrica de un cristal contiene múltiples copias de la misma

molécula. La aplicación de este método es más poderosa cuando los y los datos de difracción no han permitido definir diferencias entre las copias.

- Afinamiento usando *TLS* (Winn *et al.*, 2001), que alude a los parámetros de Traslación, Libración y movimiento de roto-traslación (en inglés “Translation”, “Libration” y “Screw”). El número de parámetros afinados usando la noción de TLS son 20 por cada “grupo TLS”. Sobre cada uno de estos grupos se aplica un afinamiento de los factores *B* anisotrópico utilizando estos 20 parámetros. Los “grupos de TLS” fueron seleccionados usando la herramienta informática *TLSMD* (Painter y Merritt, 2006) a partir de un modelo en el que todavía no se había aplicado un afinamiento anisotrópico de los factores de desplazamiento atómicos. Este método reduce el número de variables, ya que se aplican de forma estándar los 20 parámetros a cada grupo, aunque difieran en número de átomos, en lugar de los seis parámetros por átomo necesarios en un afinamiento anisotrópico completo.

1.5 Validación de la estructura

Diferentes indicadores dan una idea de la fiabilidad del modelo trazado. El *factor R* (calculado según la Ecuación 4), representa un indicador de la adecuación del modelo, mostrando la convergencia entre los datos afinados y el modelo teórico, de forma que un descenso del mismo indica un aumento en la adecuación del modelo.

$$R = \sum_{hkl} [|F_{obs}(hkl)| - |F_{calc}(hkl)|] / \sum_{hkl} |F_{obs}(hkl)|$$

Ecuación 4. Cálculo del *factor R*, que muestra la adecuación del modelo. “*F_{obs}(hkl)*” es el factor de estructura observado en las difracciones y “*F_{calc}(hkl)*” es el factor de estructura calculado a partir del modelo.

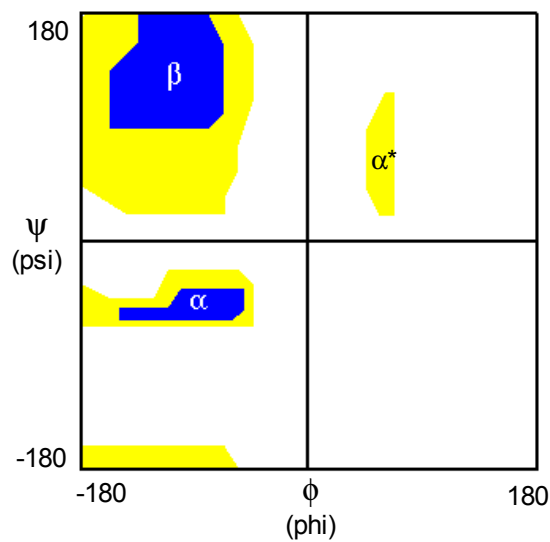
Este *factor R* marca la diferencia relativa entre los factores de estructura *F_{obs}* observados y *F_{calc}* (calculados de forma hipotética). Nunca se consigue un *factor R* = 0, y para macromoléculas se acepta como bueno un error de entre el 15 y el 25% al final del afinado.

Junto con el *factor R*, se calcula un *factor “R_{free}”* (similar al *factor R* pero calculado con observaciones que no participaron en el proceso de afinamiento).

Es necesario comprobar al final del afinado si todos los parámetros estructurales y geométricos que definen la estructura se mantienen en rangos razonables y válidos. Un ejemplo de ellos es la desviación cuadrática media (en inglés, “*r.m.s.d.*” por “root mean square deviation”) sobre los ángulos y la longitud de los enlaces. Las desviaciones en torno a los valores medios de estos parámetros dependen de la resolución de la estructura y del método de afinamiento utilizado. Un modelo afinado a una resolución aproximadamente 2 Å, debería presentar un *r.m.s.d.* de no mas de 0,01 Å para los enlaces y de 1,4 ° para los ángulos (Ref. bibliográfica 2*).

Los ángulos “phi; ϕ ” y “psi; ψ ” permitidos entre los enlaces diedros de la cadena principal $C_{i-1}-N_i-C\alpha_i-C_i$ y $N_i-C\alpha_i-C_i-N_{i+1}$ respectivamente fueron chequeados mediante un diagrama de “*Ramachandran*” (Kleywegt y Jones, 1996; Lovell *et al.*, 2003) (Figura 91). Un punto en este diagrama representa un residuo con unos determinados ángulos ϕ y ψ a las dos caras del carbono α , los cuales deben corresponder a elementos de estructura secundaria tipo α -hélice o β -lámina o bien a estructuras extendidas en bucle (en inglés “loop”) o giro (“turn”).

Figura 90. Diagrama de *Ramachandran*. En los ejes x e y se muestran respectivamente los valores para los ángulos ϕ (entre $N-C\alpha$) y ψ (entre $C-C\alpha$) para un polipéptido. Las regiones blancas corresponden a conformaciones donde los átomos están mas cercanos que la suma de sus radios de “*van der Waals*”, por lo que son estéricamente regiones energéticamente desfavorecidas para todos los aminoácidos excepto para las glicinas, que carecen de cadena lateral. Las áreas azules y amarillas son regiones permitidas, energéticamente favorables. En las regiones amarillas se sitúan residuos que mantienen distancias mas cortas entre sus átomos, cuando se usan restricciones de radios mas cortos en el cálculo. La región amarilla denominada con α^* alude a una conformación menos común de hélice alfa que torna a la izquierda (en inglés “left handed alpha-helix”).



Bibliografia

Acharya A, Rishi V, Moll J, Vinson C. Experimental identification of homodimerizing B-ZIP families in Homo sapiens. *J Struct Biol* (2006) **155**: 130-9.

Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* (2002) **58**: 1948-54.

Alam SL, Sun J, Payne M, Welch BD, Blake BK, Davis DR, Meyer HH, Emr SD, Sundquist WI. Ubiquitin interactions of NZF zinc fingers. *EMBO J* (2004) **23**: 1411-21.

Arnulf B, Villemain A, Nicot C, Mordelet E, Charneau P, Kersual J, Zermati Y, Mauviel A, Bazarbachi A, Hermine O. Human T-cell lymphotropic virus oncoprotein Tax represses TGF-beta 1 signaling in human T cells via c-Jun activation: a potential mechanism of HTLV-I leukemogenesis. *Blood* (2002) **100**: 4129-38.

Bairoch A, Apweiler R. The SWISS-PROT protein sequence database: its relevance to human molecular medical research. *J Mol Med* (1997) **75**: 312-6.

Ballard DW, Böhnlein E, Lowenthal JW, Wano Y, Franza BR, Greene WC. HTLV-I tax induces cellular proteins that activate the kappa B element in the IL-2 receptor alpha gene. *Science* (1988) **241**: 1652-5.

Banfield MJ, Barker JJ, Perry AC, Brady RL. Function from structure? The crystal structure of human phosphatidylethanolamine-binding protein suggests a role in membrane signal transduction. *Structure* (1998) **6**: 1245-54.

Bannasch D, Weis I, Schwab M. Nmi protein interacts with regions that differ between MycN and Myc and is localized in the cytoplasm of neuroblastoma cells in contrast to nuclear MycN. *Oncogene* (1999) **18**: 6810-7.

Barkett M, Gilmore TD. Control of apoptosis by Rel/NF-kappaB transcription factors. *Oncogene* (1999) **18**: 6910-24.

Barnes PJ, Karin M. Nuclear factor-kappaB: a pivotal transcription factor in chronic inflammatory diseases. *N Engl J Med* (1997) **336**: 1066-71.

Berger JC, Vander Griend D, Stadler WM, Rinker-Schaeffer C. Metastasis suppressor genes: signal transduction, cross-talk and the potential for modulating the behavior of metastatic cells. *Anticancer Drugs* (2004) **15**: 559-68.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res* (2000) **28**: 235-42.

Bernards R, Weinberg RA. A progression puzzle. *Nature* (2002) **418**: 823.

Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, Kannouche P, Wider G, Peter M, Lehmann AR, Hofmann K, Dikic I. Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science* (2005) **310**: 1821-4.

Bomar MG, Pai M, Tzeng S, Li SS, Zhou P. Structure of the ubiquitin-binding zinc finger domain of human DNA Y-polymerase eta. *EMBO Rep* (2007) **8**: 247-51.

Bonnet J, Romier C, Tora L, Devys D. Zinc-finger UBPs: regulators of deubiquitylation. *Trends Biochem Sci* (2008) **33**: 369-75.

Breslauer KJ, Frank R, Blöcker H, Marky LA. Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci U S A* (1986) **83**: 3746-50.

Brown JH, Cohen C, Parry DA. Heptad breaks in alpha-helical coiled coils: stutters and stammers. *Proteins* (1996) **26**: 134-45.

Campos-Olivas R, Sánchez R, Torres D, Blanco FJ. Backbone assignment of the 98 kDa homotrimeric yeast PCNA ring. *J Biomol NMR* (2007) **38**: 167.

Cao Z, Xiong J, Takeuchi M, Kurama T, Goeddel DV. TRAF6 is a signal transducer for interleukin-1. *Nature* (1996) **383**: 443-6.

CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* (1994) **50**: 760-3.

Chen J, Shpall RL, Meyerdierks A, Hagemeyer M, Böttger EC, Naumovski L. Interferon-inducible Myc/STAT-interacting protein Nmi associates with IFP 35 into a high molecular mass complex and inhibits proteasome-mediated degradation of IFP 35. *J Biol Chem* (2000) **275**: 36278-84.

Chen L, Mu Y, Greene WC. Acetylation of RelA at discrete sites regulates distinct nuclear functions of NF-kappaB. *EMBO J* (2002) **21**: 6539-48.

Chen L, Williams SA, Mu Y, Nakano H, Duerr JM, Buckbinder L, Greene WC. NF-kappaB RelA phosphorylation regulates RelA acetylation. *Mol Cell Biol* (2005) **25**: 7966-75.

Chen LF, Fischle W, Verdin E, Greene WC. Duration of nuclear NF-kappaB action regulated by reversible acetylation. *Science* (2001) **293**: 1653-7.

Chen ZJ, . Ubiquitin signalling in the NF-kappaB pathway. *Nat Cell Biol* (2005) **7**: 758-65.

Chin K, Chun ACS, Ching Y, Jeang K, Jin D. Human T-cell leukemia virus oncoprotein tax represses nuclear receptor-dependent transcription by targeting coactivator TAX1BP1. *Cancer Res* (2007) **67**: 1072-81.

Chu ZL, Shin YA, Yang JM, DiDonato JA, Ballard DW. IKKgamma mediates the interaction of cellular IkappaB kinases with the tax transforming protein of human T cell leukemia virus type 1. *J Biol Chem* (1999) **274**: 15297-300.

Cicek M, Fukuyama R, Welch DR, Sizemore N, Casey G. Breast cancer metastasis suppressor 1 inhibits gene expression by targeting nuclear factor-kappaB activity. *Cancer Res* (2005) **65**: 3586-95.

Cordier F, Grubisha O, Traincard F, Véron M, Delepierre M, Agou F. The zinc finger of NEMO is a functional ubiquitin-binding domain. *J Biol Chem* (2008) : .

Corpet F, . Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* (1988) **16**: 10881-90.

Cowtan K, . The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* (2006) **62**: 1002-11.

Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB3, Snoeyink J, Richardson JS, Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* (2007) **35**: W375-83.

De Valck D, Jin DY, Heyninck K, Van de Craen M, Contreras R, Fiers W, Jeang KT, Beyaert R. The zinc finger protein A20 interacts with a novel anti-apoptotic protein which is cleaved by specific caspases. *Oncogene* (1999) **18**: 4182-90.

DeWald DB, Torabinejad J, Samant RS, Johnston D, Erin N, Shope JC, Xie Y, Welch DR. Metastasis suppression by breast cancer metastasis suppressor 1 involves reduction of phosphoinositide signaling in MDA-MB-435 breast carcinoma cells. *Cancer Res* (2005) **65**: 713-7.

Doucas V, Evans RM. Human T-cell leukemia retrovirus-Tax protein is a repressor of nuclear receptor signaling. *Proc Natl Acad Sci U S A* (1999) **96**: 2633-8.

Doyon Y, Cayrou C, Ullah M, Landry A, Côté V, Selleck W, Lane WS, Tan S, Yang X, Côté J. ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Mol Cell* (2006) **21**: 51-64.

Dunn SE, Torres JV, Oh JS, Cykert DM, Barrett JC. Up-regulation of urokinase-type plasminogen activator by insulin-like growth factor-I depends upon phosphatidylinositol-3 kinase and mitogen-activated protein kinase kinase. *Cancer Res* (2001) **61**: 1367-74.

Eccles SA, Welch DR. Metastasis: recent discoveries and novel treatment strategies. *Lancet* (2007) **369**: 1742-57.

el-Sabban ME, Pauli BU. Cytoplasmic dye transfer between metastatic tumor cells and vascular endothelium. *J Cell Biol* (1991) **115**: 1375-82.

el-Sabban ME, Pauli BU. Adhesion-mediated gap junctional communication between lung-metastatic cancer cells and endothelium. *Invasion Metastasis* (1994-1995) **14**: 164-76.

Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* (2004) **60**: 2126-32.

Fattaey AR, Helin K, Dembski MS, Dyson N, Harlow E, Vuocolo GA, Hanobik MG, Haskell KM, Oliff A, Defeo-Jones D, et al.. Characterization of the retinoblastoma binding proteins RBP1 and RBP2. *Oncogene* (1993) **8**: 3149-56.

Finn RD, Tate J, Mistry J, Coghill PC, Sammut SJ, Hotz H, Ceric G, Forslund K, Eddy SR, Sonnhammer ELL, Bateman A. The Pfam protein families database. *Nucleic Acids Res* (2008) **36**: D281-8.

Franzoso G, Bours V, Park S, Tomita-Yamaguchi M, Kelly K, Siebenlist U. The candidate oncoprotein Bcl-3 is an antagonist of p50/NF-kappa B-mediated inhibition. *Nature* (1992) **359**: 339-42.

Frisch SM, Francis H. Disruption of epithelial cell-matrix interactions induces apoptosis. *J Cell Biol* (1994) **124**: 619-26.

Fu M, Wang C, Zhang X, Pestell RG. Acetylation of nuclear receptors in cellular growth and apoptosis. *Biochem Pharmacol* (2004) **68**: 1199-208.

Fujita N, Nagahashi A, Nagashima K, Rokudai S, Tsuruo T. Acceleration of apoptotic cell death after the cleavage of Bcl-XL protein by caspase-3-like proteases. *Oncogene* (1998) **17**: 1295-304.

Gachon F, Peleraux A, Thebault S, Dick J, Lemasson I, Devaux C, Mesnard JM. CREB-2, a cellular CRE-dependent transcription repressor, functions in association with Tax as an activator of the human T-cell leukemia virus type 1 promoter. *J Virol* (1998) **72**: 8332-7.

Gapuzan MR, Yufit PV, Gilmore TD. Immortalized embryonic mouse fibroblasts lacking the RelA subunit of transcription factor NF-kappaB have a malignantly transformed phenotype. *Oncogene* (2002) **21**: 2484-92.

Gilmore TD, . The Rel/NF-kappaB signal transduction pathway: introduction. *Oncogene* (1999) **18**: 6842-4.

Gilmore TD, . Introduction to NF-kappaB: players, pathways, perspectives. *Oncogene* (2006) **25**: 6680-4.

Grøndahl-Hansen J, Christensen IJ, Rosenquist C, Brünner N, Mouridsen HT, Danø K, Blichert-Toft M. High levels of urokinase-type plasminogen activator and its inhibitor

PAI-1 in cytosolic extracts of breast carcinomas are associated with poor prognosis. *Cancer Res* (1993) **53**: 2513-21.

Gudmundsson J, Barkardottir RB, Eiriksdottir G, Baldursson T, Arason A, Egilsson V, Ingvarsson S. Loss of heterozygosity at chromosome 11 in breast cancer: association of prognostic factors with genetic alterations. *Br J Cancer* (1995) **72**: 696-701.

Gupta GP, Massagué J. Cancer metastasis: building a framework. *Cell* (2006) **127**: 679-95.

Haglund K, Dikic I. Ubiquitylation and cell signaling. *EMBO J* (2005) **24**: 3353-9.

Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* (2000) **100**: 57-70.

Harhaj EW, Sun SC. IKKgamma serves as a docking subunit of the IkappaB kinase (IKK) and mediates interaction of IKK with the human T-cell leukemia virus Tax protein. *J Biol Chem* (1999) **274**: 22911-4.

Harper JW, Schulman BA. Structural complexity in ubiquitin recognition. *Cell* (2006) **124**: 1133-6.

Hayden MS, Ghosh S. Signaling to NF-kappaB. *Genes Dev* (2004) **18**: 2195-224.

Hedley BD, Vaidya KS, Phadke P, Mackenzie L, Dales DW, Postenka CO, Macdonald IC, Chambers AF. BRMS1 suppresses breast cancer metastasis in multiple experimental models of metastasis by reducing solitary cell survival and inhibiting growth initiation. *Clin Exp Metastasis* (2008) : .

Heimann R, Ferguson DJ, Hellman S. The relationship between nm23, angiogenesis, and the metastatic proclivity of node-negative breast cancer. *Cancer Res* (1998) **58**: 2766-71.

Heyninck K, Beyaert R. The cytokine-inducible zinc finger protein A20 inhibits IL-1-induced NF-kappaB activation at the level of TRAF6. *FEBS Lett* (1999) **442**: 147-50.

Hicke L, Schubert HL, Hill CP. Ubiquitin-binding domains. *Nat Rev Mol Cell Biol* (2005) **6**: 610-21.

Hicks DG, Yoder BJ, Short S, Tarr S, Prescott N, Crowe JP, Dawson AE, Budd GT, Sizemore S, Cicek M, Choueiri TK, Tubbs RR, Gaile D, Nowak N, Accavitti-Loper MA, Frost AR, Welch DR, Casey G. Loss of breast cancer metastasis suppressor 1 protein expression predicts reduced disease-free survival in subsets of breast cancer patients. *Clin Cancer Res* (2006) **12**: 6702-8.

Hoberg JE, Yeung F, Mayo MW. SMRT derepression by the IkappaB kinase alpha: a prerequisite to NF-kappaB transcription and survival. *Mol Cell* (2004) **16**: 245-55.

Hoeller D, Hecker C, Dikic I. Ubiquitin and ubiquitin-like proteins in cancer pathogenesis. *Nat Rev Cancer* (2006) **6**: 776-88.

Hofmann K, Falquet L. A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems. *Trends Biochem Sci* (2001) **26**: 347-50.

Hook SS, Orian A, Cowley SM, Eisenman RN. Histone deacetylase 6 binds polyubiquitin through its zinc finger (PAZ domain) and copurifies with deubiquitinating enzymes. *Proc Natl Acad Sci U S A* (2002) **99**: 13425-30.

Hulo N, Bairoch A, Bulliard V, Cerutti L, Cuče BA, de Castro E, Lachaize C, Langendijk-Genevaux PS, Sigrist CJA. The 20 years of PROSITE. *Nucleic Acids Res* (2008) **36**: D245-9.

Hurley JH, Lee S, Prag G. Ubiquitin-binding domains. *Biochem J* (2006) **399**: 361-72.

Hurst DR, Mehta A, Moore BP, Phadke PA, Meehan WJ, Accavitti MA, Shevde LA, Hopper JE, Xie Y, Welch DR, Samant RS. Breast cancer metastasis suppressor 1 (BRMS1) is stabilized by the Hsp90 chaperone. *Biochem Biophys Res Commun* (2006) **348**: 1429-35.

Hurst DR, Xie Y, Vaidya KS, Mehta A, Moore BP, Accavitti-Loper MA, Samant RS, Saxena R, Silveira AC, Welch DR. Alterations of BRMS1-ARID4A interaction modify gene expression but still suppress metastasis in human breast cancer cells. *J Biol Chem* (2008) **283**: 7438-44.

Hutti JE, Turk BE, Asara JM, Ma A, Cantley LC, Abbott DW. IκappaB kinase beta phosphorylates the K63 deubiquitinase A20 to cause feedback inhibition of the NF-kappaB pathway. *Mol Cell Biol* (2007) **27**: 7451-61.

Iha H, Peloponese J, Verstrepen L, Zapart G, Ikeda F, Smith CD, Starost MF, Yedavalli V, Heyninck K, Dikic I, Beyaert R, Jeang K. Inflammatory cardiac valvulitis in TAX1BP1-deficient mice through selective NF-kappaB activation. *EMBO J* (2008) **27**: 629-41.

Ito A, Katoh F, Kataoka TR, Okada M, Tsubota N, Asada H, Yoshikawa K, Maeda S, Kitamura Y, Yamasaki H, Nojima H. A role for heterologous gap junctions between melanoma and endothelial cells in metastasis. *J Clin Invest* (2000) **105**: 1189-97.

Jäättelä M, Mouritzen H, Elling F, Bastholm L. A20 zinc finger protein inhibits TNF and IL-1 signaling. *J Immunol* (1996) **156**: 1166-73.

Jänicke F, Schmitt M, Ulm K, Gössner W, Graeff H. Urokinase-type plasminogen activator antigen and early relapse in breast cancer. *Lancet* (1989) **2**: 1049.

Jeang KT, Widen SG, Semmes OJ4, Wilson SH. HTLV-I trans-activator protein, tax, is a trans-repressor of the human beta-polymerase gene. *Science* (1990) **247**: 1082-4.

Jin DY, Giordano V, Kibler KV, Nakano H, Jeang KT. Role of adapter function in oncoprotein-mediated activation of NF-kappaB. Human T-cell leukemia virus type I Tax interacts directly with IkappaB kinase gamma. *J Biol Chem* (1999) **274**: 17402-5.

Jones DT, . Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* (1999) **292**: 195-202.

Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Córdón-Cardo C, Guise TA, Massagué J. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* (2003) **3**: 537-49.

Kantardjieff KA, Rupp B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci* (2003) **12**: 1865-71.

Karashima T, Sweeney P, Kamat A, Huang S, Kim SJ, Bar-Eli M, McConkey DJ, Dinney CPN. Nuclear factor-kappaB mediates angiogenesis and metastasis of human bladder cancer through the regulation of interleukin-8. *Clin Cancer Res* (2003) **9**: 2786-97.

Karin M, . Nuclear factor-kappaB in cancer development and progression. *Nature* (2006) **441**: 431-6.

Kauffman EC, Robinson VL, Stadler WM, Sokoloff MH, Rinker-Schaeffer CW. Metastasis suppression: the evolving role of metastasis suppressor genes for regulating cancer cell growth at the secondary site. *J Urol* (2003) **169**: 1122-33.

Kibler KV, Jeang KT. CREB/ATF-dependent repression of cyclin a by human T-cell leukemia virus type 1 Tax protein. *J Virol* (2001) **75**: 2161-73.

Kiernan R, Brès V, Ng RWM, Coudart M, El Messaoudi S, Sardet C, Jin D, Emiliani S, Benkirane M. Post-activation turn-off of NF-kappa B-dependent transcription is regulated by acetylation of p65. *J Biol Chem* (2003) **278**: 2758-66.

Kim J, Yu W, Kovalski K, Ossowski L. Requirement for specific proteases in cancer cell intravasation as revealed by a novel semiquantitative PCR-based assay. *Cell* (1998) **94**: 353-62.

Kim JH, Li H, Stallcup MR. CoCoA, a nuclear receptor coactivator which acts through an N-terminal activation domain of p160 coactivators. *Mol Cell* (2003) **12**: 1537-49.

Kleywegt GJ, Jones TA. Phi/psi-chology: Ramachandran revisited. *Structure* (1996) **4**: 1395-400.

Krikos A, Laherty CD, Dixit VM. Transcriptional activation of the tumor necrosis factor alpha-inducible zinc finger protein, A20, is mediated by kappa B elements. *J Biol Chem* (1992) **267**: 17971-6.

Kriminski S, Kazmierczak M, Thorne RE. Heat transfer from protein crystals: implications for flash-cooling and X-ray beam heating. *Acta Crystallogr D Biol Crystallogr* (2003) **59**: 697-708.

Lai A, Lee JM, Yang WM, DeCaprio JA, Kaelin WGJ, Seto E, Branton PE. RBP1 recruits both histone deacetylase-dependent and -independent repression activities to retinoblastoma family proteins. *Mol Cell Biol* (1999a) **19**: 6632-41.

Lai A, Marcellus RC, Corbeil HB, Branton PE. RBP1 induces growth arrest by repression of E2F-dependent transcription. *Oncogene* (1999b) **18**: 2091-100.

Langley RR, Fidler IJ. Tumor cell-organ microenvironment interactions in the pathogenesis of cancer metastasis. *Endocr Rev* (2007) **28**: 297-321.

Lee EG, Boone DL, Chai S, Libby SL, Chien M, Lodolce JP, Ma A. Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. *Science* (2000) **289**: 2350-4.

Leung K, Nabel GJ. HTLV-1 transactivator induces interleukin-2 receptor expression through an NF-kappa B-like factor. *Nature* (1988) **333**: 776-8.

Liebel U, Kindler B, Pepperkok R. 'Harvester': a fast meta search engine of human protein resources. *Bioinformatics* (2004) **20**: 1962-3.

Lin A, Karin M. NF-kappaB in cancer: a marked target. *Semin Cancer Biol* (2003) **13**: 107-14.

Ling L, Goeddel DV. T6BP, a TRAF6-interacting protein involved in IL-1 signaling. *Proc Natl Acad Sci U S A* (2000) **97**: 9567-72.

Liotta LA, Goldfarb RH, Brundage R, Siegal GP, Terranova V, Garbisa S. Effect of plasminogen activator (urokinase), plasmin, and thrombin on glycoprotein and collagenous components of basement membrane. *Cancer Res* (1981) **41**: 4629-36.

Liu Y, Smith PW, Jones DR. Breast cancer metastasis suppressor 1 functions as a corepressor by enhancing histone deacetylase 1-mediated deacetylation of RelA/p65 and promoting apoptosis. *Mol Cell Biol* (2006) **26**: 8683-96.

Locke D, . Gap junctions in normal and neoplastic mammary gland. *J Pathol* (1998) **186**: 343-9.

Locke D, Perusinghe N, Newman T, Jayatilake H, Evans WH, Monaghan P. Developmental expression and assembly of connexins into homomeric and heteromeric gap junction hemichannels in the mouse mammary gland. *J Cell Physiol* (2000) **183**: 228-37.

Look M, van Putten W, Duffy M, Harbeck N, Christensen IJ, Thomssen C, Kates R, Spyrtos F, Fernö M, Eppenberger-Castori S, Fred Sweep CGJ, Ulm K, Peyrat J, Martin P, Magdelenat H, Brünner N, Duggan C, Lisboa BW, Bendahl P, Quillien V, Daver A, Ricolleau G, Meijer-van Gelder M, Manders P, Edward Fiets W, Blankenstein M, Broët P, Romain S, Daxenbichler G, Windbichler G, Cufer T, Borstnar S, Kueng W, Beex L, Klijn J, O'Higgins N, Eppenberger U, Jänicke F, Schmitt M, Foekens J. Pooled analysis of prognostic impact of uPA and PAI-1 in breast cancer patients. *Thromb Haemost* (2003) **90**: 538-48.

Lovell SC, Davis IW, Arendall WB3, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* (2003) **50**: 437-50.

Lupas A, . Coiled coils: new structures and new functions. *Trends Biochem Sci* (1996) **21**: 375-82.

Lupas A, Van Dyke M, Stock J. Predicting coiled coils from protein sequences. *Science* (1991) **252**: 1162-4.

Mason JM, Arndt KM. Coiled coil domains: stability, specificity, and biological implications. *Chembiochem* (2004) **5**: 170-6.

Matsuoka M, . Human T-cell leukemia virus type I (HTLV-I) infection and the onset of adult T-cell leukemia (ATL). *Retrovirology* (2005) **2**: 27.

Matthews BW, . Solvent content of protein crystals. *J Mol Biol* (1968) **33**: 491-7.

Meehan WJ, Samant RS, Hopper JE, Carrozza MJ, Shevde LA, Workman JL, Eckert KA, Verderame MF, Welch DR. Breast cancer metastasis suppressor 1 (BRMS1) forms complexes with retinoblastoma-binding protein 1 (RBP1) and the mSin3 histone deacetylase complex and represses transcription. *J Biol Chem* (2004) **279**: 1562-9.

Meehan WJ, Welch DR. Breast cancer metastasis suppressor 1: update. *Clin Exp Metastasis* (2003) **20**: 45-50.

Metge BJ, Frost AR, King JA, Dyess DL, Welch DR, Samant RS, Shevde LA. Epigenetic silencing contributes to the loss of BRMS1 expression in breast cancer. *Clin Exp Metastasis* (2008) : .

Minn AJ, Kang Y, Serganova I, Gupta GP, Giri DD, Doubrovin M, Ponomarev V, Gerald WL, Blasberg R, Massagué J. Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. *J Clin Invest* (2005) **115**: 44-55.

Mishra GR, Suresh M, Kumaran K, Kannabiran N, Suresh S, Bala P, Shivakumar K, Anuradha N, Reddy R, Raghavan TM, Menon S, Hanumanthu G, Gupta M, Upendran S, Gupta S, Mahesh M, Jacob B, Mathew P, Chatterjee P, Arun KS, Sharma S, Chandrika KN, Deshpande N, Palvankar K, Raghavnath R, Krishnakanth R, Karathia H, Rekha B, Nayak R, Vishnupriya G, Kumar HGM, Nagini M, Kumar GSS, Jose R, Deepthi P, Mohan SS, Gandhi TKB, Harsha HC, Deshpande KS, Sarker M, Prasad TSK, Pandey A. Human protein reference database--2006 update. *Nucleic Acids Res* (2006) **34**: D411-4.

Mossessova E, Lima CD. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell* (2000) **5**: 865-76.

Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* (1997) **53**: 240-55.

Nakai K, Horton P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* (1999) **24**: 34-6.

Newman JRS, Keating AE. Comprehensive identification of human bZIP interactions with coiled-coil arrays. *Science* (2003) **300**: 2097-101.

Nguyen DX, Massagué J. Genetic determinants of cancer metastasis. *Nat Rev Genet* (2007) **8**: 341-52.

Nicolson GL, Dulski KM, Trosko JE. Loss of intercellular junctional communication correlates with metastatic potential in mammary adenocarcinoma cells. *Proc Natl Acad Sci U S A* (1988) **85**: 473-6.

Nicot C, Mahieux R, Pise-Masison C, Brady J, Gessain A, Yamaoka S, Franchini G. Human T-cell lymphotropic virus type 1 Tax represses c-Myb-dependent transcription through activation of the NF-kappaB pathway and modulation of coactivator usage. *Mol Cell Biol* (2001) **21**: 7391-402.

Nikolaev AY, Papanikolaou NA, Li M, Qin J, Gu W. Identification of a novel BRMS1-homologue protein p40 as a component of the mSin3A/p33(ING1b)/HDAC1 deacetylase complex. *Biochem Biophys Res Commun* (2004) **323**: 1216-22.

Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* (2006) **62**: 439-50.

Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. *Nat Rev Cancer* (2004) **4**: 448-56.

Penengo L, Mapelli M, Murachelli AG, Confalonieri S, Magri L, Musacchio A, Di Fiore PP, Polo S, Schneider TR. Crystal structure of the ubiquitin binding domains of rabex-5 reveals two modes of interaction with ubiquitin. *Cell* (2006) **124**: 1183-95.

Perkins ND, . Post-translational modifications regulating the activity and function of the nuclear factor kappa B pathway. *Oncogene* (2006) **25**: 6717-30.

Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* (1999) **6**: 458-63.

Phadke PA, Vaidya KS, Nash KT, Hurst DR, Welch DR. BRMS1 suppresses breast cancer experimental metastasis to multiple organs by inhibiting several steps of the metastatic process. *Am J Pathol* (2008) **172**: 809-17.

Phillips KK, Welch DR, Miele ME, Lee JH, Wei LL, Weissman BE. Suppression of MDA-MB-435 breast carcinoma cell metastasis following the introduction of human chromosome 11. *Cancer Res* (1996) **56**: 1222-7.

Pickart CM, Fushman D. Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol* (2004) **8**: 610-6.

Pise-Masison CA, Radonovich M, Sakaguchi K, Appella E, Brady JN. Phosphorylation of p53: a novel pathway for p53 inactivation in human T-cell lymphotropic virus type 1-transformed cells. *J Virol* (1998) **72**: 6348-55.

Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* (2007) **35**: D61-5.

Quigley JP, . Phorbol ester-induced morphological changes in transformed chick fibroblasts: evidence for direct catalytic involvement of plasminogen activator. *Cell* (1979) **17**: 131-41.

Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: integrating information about genes, proteins and diseases. *Trends Genet* (1997) **13**: 163.

Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: a novel functional genomics compendium with automated data mining and query reformulation support. *Bioinformatics* (1998) **14**: 656-64.

Reyes-Turcu FE, Horton JR, Mullally JE, Heroux A, Cheng X, Wilkinson KD. The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell* (2006) **124**: 1197-208.

Rinker-Schaeffer CW, O'Keefe JP, Welch DR, Theodorescu D. Metastasis suppressor proteins: discovery, molecular mechanisms, and clinical application. *Clin Cancer Res* (2006) **12**: 3882-9.

Riou P, Bex F, Gazzolo L. The human T cell leukemia/lymphotropic virus type 1 Tax protein represses MyoD-dependent transcription by inhibiting MyoD-binding to the KIX domain of p300. A potential mechanism for Tax-mediated repression of the transcriptional activity of basic helix-loop-helix factors. *J Biol Chem* (2000) **275**: 10551-60.

Rivera J, Megias D, Bravo J. Proteomics-based strategy to delineate the molecular mechanisms of the metastasis suppressor gene BRMS1. *J Proteome Res* (2007) **6**: 4006-18.

Safran M, Solomon I, Shmueli O, Lapidot M, Shen-Orr S, Adato A, Ben-Dor U, Esterman N, Rosen N, Peter I, Olender T, Chalifa-Caspi V, Lancet D. GeneCards 2002: towards a complete, object-oriented, human gene compendium. *Bioinformatics* (2002) **18**: 1542-3.

Salerno WJ, Seaver SM, Armstrong BR, Radhakrishnan I. MONSTER: inferring non-covalent interactions in macromolecular structures from atomic coordinate data. *Nucleic Acids Res* (2004) **32**: W566-8.

Samant RS, Clark DW, Fillmore RA, Cicek M, Metge BJ, Chandramouli KH, Chambers AF, Casey G, Welch DR, Shevde LA. Breast cancer metastasis suppressor 1 (BRMS1) inhibits osteopontin transcription by abrogating NF-kappaB activation. *Mol Cancer* (2007) **6**: 6.

Samant RS, Debies MT, Shevde LA, Verderame MF, Welch DR. Identification and characterization of the murine ortholog (brms1) of breast-cancer metastasis suppressor 1 (BRMS1). *Int J Cancer* (2002) **97**: 15-20.

Samant RS, Seraj MJ, Saunders MM, Sakamaki TS, Shevde LA, Harms JF, Leonard TO, Goldberg SF, Budgeon L, Meehan WJ, Winter CR, Christensen ND, Verderame MF, Donahue HJ, Welch DR. Analysis of mechanisms underlying BRMS1 suppression of metastasis. *Clin Exp Metastasis* (2000) **18**: 683-93.

Saunders MM, Seraj MJ, Li Z, Zhou Z, Winter CR, Welch DR, Donahue HJ. Breast cancer metastatic potential correlates with a breakdown in homospecific and heterospecific gap junctional intercellular communication. *Cancer Res* (2001) **61**: 1765-7.

Schindl M, Birner P, Breitenacker G, Oberhuber G. Downregulation of KAI1 metastasis suppressor protein is associated with a dismal prognosis in epithelial ovarian cancer. *Gynecol Oncol* (2001) **83**: 244-8.

Seigneurin-Berny D, Verdel A, Curtet S, Lemerrier C, Garin J, Rousseaux S, Khochbin S. Identification of components of the murine histone deacetylase 6 complex: link between acetylation and ubiquitination signaling pathways. *Mol Cell Biol* (2001) **21**: 8035-44.

Seraj MJ, Harding MA, Gildea JJ, Welch DR, Theodorescu D. The relationship of BRMS1 and RhoGDI2 gene expression to metastatic potential in lineage related human bladder cancer cell lines. *Clin Exp Metastasis* (2000a) **18**: 519-25.

Seraj MJ, Samant RS, Verderame MF, Welch DR. Functional evidence for a novel human breast carcinoma metastasis suppressor, BRMS1, encoded at chromosome 11q13. *Cancer Res* (2000b) **60**: 2764-9.

Sheldrick GM, . A short history of SHELX. *Acta Crystallogr A* (2008) **64**: 112-22.

Shembade N, Harhaj NS, Liebl DJ, Harhaj EW. Essential role for TAX1BP1 in the termination of TNF-alpha-, IL-1- and LPS-mediated NF-kappaB and JNK signaling. *EMBO J* (2007) **26**: 3910-22.

Shembade N, Harhaj NS, Parvatiyar K, Copeland NG, Jenkins NA, Matesic LE, Harhaj EW. The E3 ligase Itch negatively regulates inflammatory signaling pathways by controlling the function of the ubiquitin-editing enzyme A20. *Nat Immunol* (2008) **9**: 254-62.

Shevde LA, Samant RS, Goldberg SF, Sikaneta T, Alessandrini A, Donahue HJ, Mauger DT, Welch DR. Suppression of human melanoma metastasis by the metastasis suppressor gene, BRMS1. *Exp Cell Res* (2002) **273**: 229-39.

Shevde LA, Welch DR. Metastasis suppressor pathways--an evolving paradigm. *Cancer Lett* (2003) **198**: 1-20.

Sliva D, Rizzo MT, English D. Phosphatidylinositol 3-kinase and NF-kappaB regulate motility of invasive MDA-MB-231 human breast cancer cells by the secretion of urokinase-type plasminogen activator. *J Biol Chem* (2002) **277**: 3150-7.

Smid M, Wang Y, Klijn JGM, Sieuwerts AM, Zhang Y, Atkins D, Martens JWM, Foekens JA. Genes associated with breast cancer metastatic to bone. *J Clin Oncol* (2006) **24**: 2261-7.

Spínola-Amilibia M, Rivera J, Bravo J. Crystallization and preliminary X-ray diffraction analysis of a breast cancer metastasis suppressor 1 predicted coiled-coil region. *Acta Crystallogr Sect F Struct Biol Cryst Commun* (2008) **64**: 1156-8.

Stafford LJ, Vaidya KS, Welch DR. Metastasis suppressors genes in cancer. *Int J Biochem Cell Biol* (2008) **40**: 874-91.

Steeg PS, . Metastasis suppressors alter the signal transduction of cancer cells. *Nat Rev Cancer* (2003) **3**: 55-63.

Steeg PS, Bevilacqua G, Kopper L, Thorgerirsson UP, Talmadge JE, Liotta LA, Sobel ME. Evidence for a novel gene associated with low tumor metastatic potential. *J Natl Cancer Inst* (1988) **80**: 200-4.

Steeg PS, Theodorescu D. Metastasis: a therapeutic target for cancer. *Nat Clin Pract Oncol* (2008) **5**: 206-19.

Strelkov SV, Burkhard P. Analysis of alpha-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. *J Struct Biol* (2002) **137**: 54-64.

Sudol M, Chen HI, Bougeret C, Einbond A, Bork P. Characterization of a novel protein-binding module--the WW domain. *FEBS Lett* (1995) **369**: 67-71.

Swanson KA, Kang RS, Stamenova SD, Hicke L, Radhakrishnan I. Solution structure of Vps27 UIM-ubiquitin complex important for endosomal sorting and receptor downregulation. *EMBO J* (2003) **22**: 4597-606.

Terwilliger TC, . Automated structure solution, density modification and model building. *Acta Crystallogr D Biol Crystallogr* (2002) **58**: 1937-40.

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* (2008) **64**: 61-9.

UniProt. The Universal Protein Resource (UniProt). *Nucleic Acids Res* (2007) **35**: D193-7.

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* (2002) **415**: 530-6.

Van PL, Yim KW, Jin DY, Dapolito G, Kurimasa A, Jeang KT. Genetic evidence of a role for ATM in functional interaction between human T-cell leukemia virus type 1 Tax and p53. *J Virol* (2001) **75**: 396-407.

Vinson C, Acharya A, Taparowsky EJ. Deciphering B-ZIP transcription factor interactions in vitro and in vivo. *Biochim Biophys Acta* (2006) **1759**: 4-12.

Vonrhein C, Blanc E, Roversi P, Bricogne G. Automated structure solution with autoSHARP. *Methods Mol Biol* (2007) **364**: 215-30.

Welch DR, Steeg PS, Rinker-Schaeffer CW. Molecular biology of breast cancer metastasis. Genetic regulation of human breast carcinoma metastasis. *Breast Cancer Res* (2000) **2**: 408-16.

- Wertz IE, O'Rourke KM, Zhou H, Eby M, Aravind L, Seshagiri S, Wu P, Wiesmann C, Baker R, Boone DL, Ma A, Koonin EV, Dixit VM. De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. *Nature* (2004) **430**: 694-9.
- Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* (2001) **57**: 122-33.
- Wolf E, Kim PS, Berger B. MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci* (1997) **6**: 1179-89.
- Wu K, Bottazzi ME, de la Fuente C, Deng L, Gitlin SD, Maddukuri A, Dadgar S, Li H, Vertes A, Pumfery A, Kashanchi F. Protein profile of tax-associated complexes. *J Biol Chem* (2004) **279**: 495-508.
- Wullaert A, Heyninck K, Janssens S, Beyaert R. Ubiquitin: tool and target for intracellular NF-kappaB inhibitors. *Trends Immunol* (2006) **27**: 533-40.
- Yang CK, Kim JH, Li H, Stallcup MR. Differential use of functional domains by coiled-coil coactivator in its synergistic coactivator function with beta-catenin or GRIP1. *J Biol Chem* (2006) **281**: 3389-97.
- Yang J, Zhang B, Lin Y, Yang Y, Liu X, Lu F. Breast cancer metastasis suppressor 1 inhibits SDF-1alpha-induced migration of non-small cell lung cancer by decreasing CXCR4 expression. *Cancer Lett* (2008) : .
- Yoshida M, . Multiple viral strategies of HTLV-1 for dysregulation of cell growth control. *Annu Rev Immunol* (2001) **19**: 475-96.
- Zemzoum I, Kates RE, Ross JS, Dettmar P, Dutta M, Henrichs C, Yurdseven S, Höfler H, Kiechle M, Schmitt M, Harbeck N. Invasion factors uPA/PAI-1 and HER2 status provide independent and complementary information on patient outcome in node-negative breast cancer. *J Clin Oncol* (2003) **21**: 1022-8.
- Zhang S, Lin Q, DI W. Suppression of human ovarian carcinoma metastasis by the metastasis-suppressor gene, BRMS1. *Int J Gynecol Cancer* (2006) **16**: 522-31.
- Zhu M, John S, Berg M, Leonard WJ. Functional association of Nmi with Stat5 and Stat1 in IL-2- and IFNgamma-mediated signaling. *Cell* (1999) **96**: 121-30.

Referencias

Referencias bibliográficas

- *1 Innis, M. A., Gelfand, D. H. 1990. *Optimization of PCRs*, pp. 3–20. In: M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (eds.), *PCR protocols, a guide to methods and applications*. Academic Press, San Diego, CA.
- *2 Gale Rhodes. 1993. *Crystallography. Made Crystal Clear*, Second edition. Academic Press, San Diego, CA. London, UK.
- *3 Sylvie Doublé. 2007. *Methods in Molecular Biology*, vol. 364: *Macromolecular Crystallography Protocols*, Volume 2, Structure Determination. Humana Press, Totowa, NJ.
- *4 Sambrook, J.; Fritsch, EF.; Maniatis, T. 1989. *Molecular Cloning - a laboratory manual*. Second edition. Cold Spring Harbor Laboratory Press, New York.
- *5 Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. 2005. *Protein Identification and Analysis Tools on the ExPASy Server*, pp. 571-607. (In) John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press.
- *6 Ausubel, F. M. et al. 1987. *Current Protocols in Molecular Biology*, Second edition. New York: Greene Publishing Associates and Wiley.

Referencias informáticas

- ¹ **PSORT II** (Nakai y Horton, 1999): <http://psort.ims.u-tokyo.ac.jp/form2.html>
- ² **PROSITE** (Hulo et al., 2008): <http://www.expasy.ch/prosite/>
- ³ **ExPASy** (Swiss Institute of Bioinformatics (SIB)): <http://www.expasy.org/>
- ⁴ **UniProtKB/Swiss-Prot** (Bairoch y Apweiler, 1997; UniProt, 2007) (Swiss Institute of Bioinformatics (SIB)): <http://expasy.org/sprot/>
- ⁵ **BMERC**: <http://bmerc-www.bu.edu/>
- ⁶ **Bioinformatic Harvester (EMBL)** (Liebel et al., 2004): <http://harvester.embl.de/>
(Actualmente <http://harvester.fzk.de>)
- ⁷ **NCBI** (Pruitt et al., 2007): <http://www.ncbi.nlm.nih.gov/RefSeq/>
- ⁸ **GeneCards** (Rebhan et al., 1997; Rebhan et al., 1998; Safran et al., 2002): <http://www.genecards.org/>
- ⁹ **Human Protein Reference Database (hprd)** (Mishra et al., 2006): <http://www.hprd.org/>
- ¹⁰ **PDB** (Berman et al., 2000): <http://www.rcsb.org/pdb/home/home.do>

- ¹¹ **PSIPRED** (Jones, 1999): <http://bioinf.cs.ucl.ac.uk/psipred/>
- ¹² **HomoloGene (NCBI)**: <http://www.ncbi.nlm.nih.gov/sites/entrez?db=homologene>
- ¹³ **MultAlin** (Corpet, 1988): <http://bioinfo.genotoul.fr/multalin/multalin.html>
- ¹⁴ **Coiled-coil prediction** (Lupas *et al.*, 1991): <http://www.russell.embl-heidelberg.de/cgi-bin/coils-svr.pl>
- ¹⁵ **Multicoil** (Wolf *et al.*, 1997): <http://groups.csail.mit.edu/cb/multicoil/cgi-bin/multicoil.cgi>
- ¹⁶ **pbil** (Lupas *et al.*, 1991): http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_lupas.html
- ¹⁷ **Proligo** (Breslauer *et al.*, 1986): <http://www.proligo.com/Calculation/calculation.html>
- ¹⁸ **ProtParam (Expasy)** (Ref. bibliográfica 5*): <http://expasy.org/tools/protparam.html>
- ¹⁹ **Pfam** (Finn *et al.*, 2008): <http://pfam.sanger.ac.uk/search>

Casas comerciales

- ^A **MJ Research**: <http://www.labtrade.com/mjr/mjr.htm>
- ^B **Promega**: <http://www.promega.com/Default.asp>
- ^C **Fermentas**: <http://www.fermentas.com/>
- ^D **Proligo**: <http://www.proligo.com/>
- ^E **BioRad**: www.bio-rad.com/
- ^F **Takara**: <http://www.takara-bio.eu/>
- ^G **Qiagen-Nextal**: <http://www1.qiagen.com/>
- ^H **Novagen**: <http://www.emdbiosciences.com/html/NVG/home.html>
- ^I **GE Healthcare-Amersham Biosciences**: <http://www.gehealthcare.com/eses/>
- ^J **Roche**: <http://www.roche.com/home.html>
- ^K **Genomed**: <http://www.genomed-dna.com/>
- ^L **Invitrogen**: <http://www.invitrogen.com/>
- ^M **Eppendorf**: <http://www.eppendorf.es>
- ^N **Promega**: <http://www.promega.com/>

- N BostonBiochem:** <http://www.bostonbiochem.com/>
- O New Brunswick Scientific:** <http://www.nbsc.com/>
- P Beckman Coulter:** <http://www.beckmancoulter.com/>
- Q Fluka-Sigma-Aldrich:**
http://www.sigmaaldrich.com/Brands/Fluka___Riedel_Home.html
- R Fisher Scientific Bioblock:** <http://www.bioblock.com/pages/home.asp>
- S Pierce:** <http://www.piercenet.com/>
- T Millipore:** <http://www.millipore.com>
- U Emerald Biosystems:** <http://emeraldbiosystems.com/dnn/entity/tabid/103/c-26-wizard.aspx>
- V Hampton Research:** <http://www.hamptonresearch.com/>
- W Jena Bioscience:** <http://www.jenabioscience.com/>
- X Genomic solutions:** <http://www.genomicsolutions.com/>
- Y Greiner:**
- Z Costar:**
- AB Innovaplate:**
- AC Tecan:** <http://www.tecan.com>
- AD Molecular Dimensions:** <http://www.moleculardimensions.com/>
- AE NanoDrop (Thermo scientific):** <http://www.nanodrop.com/>

Tablas y Figuras

Tablas

Tabla I.	Recopilación de proteínas supresoras de metástasis y proceso celular en el que están involucradas	26
Tabla II.	Materiales usados en los clonajes	54
Tabla III.	Tampones usados en las purificaciones de las construcciones con cola de histidinas	59
Tabla IV.	Tampones usados en las purificaciones de las construcciones fusionadas a GST	61
Tabla V.	Líneas del sincrotrón de Grenoble utilizadas	70
Tabla VI.	Isoformas descritas para las proteínas BRMS1 y Tax1BP1	77
Tabla VII.	Resumen de los resultados obtenidos para las construcciones de las proteínas humanas BRMS1 y NMI	84
Tabla VIII.	Parámetros usados en la recolección de los juegos de datos a alta y baja resolución de un cristal de BRMS151-84	94
Tabla IX.	Derivados de átomos pesados de platino, oro y mercurio, seleccionados para impregnar los cristales de BRMS151-84	96
Tabla X.	Parámetros usados en la recolección del conjunto de datos de difracción de un cristal de BRMS151-84 derivado de mercurio	97
Tabla XI.	Datos estadísticos cristalográficos obtenidos tras el procesamiento y simplificación de los datos de difracción recogidos de los cristales nativo y derivado de mercurio de la proteína BRMS151-84	99
Tabla XII.	Posibles contenidos de moléculas y solvente en la unidad asimétrica (UA) para el cristal nativo de BRMS151-84	100
Tabla XIII.	Datos estadísticos cristalográficos obtenidos tras el proceso de afinamiento del modelo obtenido de BRMS151-84	103
Tabla XIV.	Residuos modelados de las 8 cadenas (A-H) de la unidad asimétrica del cristal de BRMS151-84	104
Tabla XV.	Parámetros usados para la recolección del juego de datos de difracción de un cristal de BRMS151-98	113
Tabla XVI.	Resultados estadísticos cristalográficos obtenidos tras el procesamiento y simplificación de los datos de difracción recogidos del cristal de BRMS151-98	114
Tabla XVII.	Parámetros usados en la recolección del conjunto de datos de difracción de un cristal de UBZ1+2	138
Tabla XVIII.	Datos estadísticos cristalográficos obtenidos tras el procesamiento y simplificación de los datos de difracción recogidos del cristal de UBZ1+2	139

Tabla XIX.	Coeficientes de Matthews y probabilidades normalizadas para cada posible número de moléculas en la unidad asimétrica (UA) y porcentaje de solvente para el cristal procesado de UBZ1+2	140
Tabla XX.	Datos estadísticos cristalográficos obtenidos tras el proceso de afinamiento del modelo final obtenido de UBZ1+2	142
Tabla XXI.	Residuos modelados para las 11 cadenas (A-K) que componen la unidad asimétrica del cristal de UBZ1+2	143

Figuras

Figura 1.	Etapas en el proceso de metástasis	23
Figura 2.	Modelo hipotético de interferencia entre productos de genes supresores de metástasis en vías de señalización Rac, Src y Raf-1	27
Figura 3.	Representación esquemática de la predicción de dominios de la proteína BRMS1 humana	29
Figura 4.	Representación esquemática de la predicción de dominios de la proteína Tax1BP1 humana	34
Figura 5.	Alineamiento de secuencia de los dominios estructurados en “dedos de zinc” de los homólogos de Tax1BP1 humano en diferentes especies	34
Figura 6.	Rutas de señalización del factor de transcripción NFκB	39
Figura 7.	Representación de la regulación inhibitoria de BRMS1 y Tax1BP1 sobre diferentes etapas de la ruta de activación de NFκB	44
Figura 8.	Material usado para el montaje y la conservación de los cristales	69
Figura 9.	Alineamiento basado en la secuencia aminoacídica de la proteína BRMS1 entre diferentes especies	81
Figura 10.	Representación esquemática de la predicción de dominios de la proteína NMI humana	82
Figura 11.	Secuencia originada del clonaje del fragmento comprendiendo los aminoácidos 51 al 84 de BRMS1 antes y después de su proteólisis	86
Figura 12.	Expresión de His-SMT3-BRMS151-84 en cepa de E.coli	86
Figura 13.	Cromatograma y muestras resultantes de la columna de afinidad a histidinas en una purificación de His-SMT3-BRMS151-84	87
Figura 14.	Muestras resultantes de la proteólisis de His-SMT3-BRMS151-84	88
Figura 15.	Cromatografía de exclusión por tamaño molecular del polipéptido BRMS151-84	89
Figura 16.	Espectroscopía de Dicroísmo Circular del polipéptido BRMS151-84	90
Figura 17.	Imágenes de cristales iniciales de BRMS151-84 obtenidos mediante la técnica de difusión de vapor en gota sentada	91

Figura 18.	Cristales de BRMS151-84 optimizados obtenidos mediante la técnica de difusión de vapor en gota colgante	92
Figura 19.	Cristales de BRMS151-84 crecidos a lo largo de la ruta trazada con un pelo de gato, mediante la técnica de sembrado de centros de nucleación o “microseeding”.	93
Figura 20.	Imágenes representativas de espectros de difracción de los juegos de datos recogidos a alta y a baja resolución de un cristal de BRMS151-84	94
Figura 21.	Electroforesis en gel de poliacrilamida en condiciones nativas de muestras conteniendo proteína BRMS151-84 incubada con 24 derivados de átomos pesados	95
Figura 22.	Ejemplo de cristales de BRMS151-84 mejorados usados para la obtención de un cristal derivado de átomo pesado	97
Figura 23.	Patrón de difracción de un cristal de BRMS151-84 derivado con nitrato de mercurio	98
Figura 24.	Ejemplo de sección de Harker de un mapa de Patterson de diferencias para un cristal de BRMS151-84 derivado de mercurio	100
Figura 25.	Mapa de densidad electrónica experimental inicial, calculado por Coot tras asignar las fases usando la técnica SIRAS, para la resolución de la estructura BRMS151-84	101
Figura 27.	Estrategia utilizada para resolver la estructura de BRMS151-84 mediante SIRAS	102
Figura 28.	Unidad asimétrica conteniendo 8 moléculas (A-H) de BRMS151-84	104
Figura 29.	Puente disulfuro establecido entre los residuos de cisteína 10 de las cadenas C y H contenidas en la unidad asimétrica del cristal de BRMS151-84	105
Figura 30.	Contactos mantenidos entre los residuos de las dos cadenas A-Asim2 de BRMS151-84, analizados con el programa Monster	106
Figura 31.	Estructura de las dos cadenas A y Asim2	107
Figura 32.	Diagrama de Ramachandran conteniendo los 192 residuos de la unidad asimétrica del cristal de BRMS151-84	108
Figura 33.	Validación del modelo, según el ajuste de los residuos de las ocho cadenas de la unidad asimétrica de BRMS151-84 a la densidad electrónica, y a la geometría	108
Figura 34.	Representación de la variación atómica de los factores B de las 8 cadenas de BRMS151-84 que forman la unidad asimétrica	109
Figura 35.	Expresión de His-SMT3-BRMS151-98 en cepa de E.coli	110
Figura 36.	Cromatografía de exclusión por tamaño molecular de BRMS151-98 libre de His-SMT3	111
Figura 37.	Dispersión Dinámica de la Luz de BRMS151-98	112
Figura 38.	Gel de poliacrilamida mostrando el elevado grado de pureza de la proteína final BRMS151-98	112

Figura 39.	Cristales de BRMS151-98 obtenidos mediante la técnica de difusión de vapor	113
Figura 40.	Imagen representativa del patrón de difracción obtenida en la recolección de los datos del cristal de BRMS151-98	113
Figura 41.	Expresión de His-SMT3-BRMS151-98 con selenometionina en la cepa Rosetta(DE3)	115
Figura 42.	Difracción de un cristal de SeMet-BRMS151-98	116
Figura 43.	Expresión de His-SMT3-NMI16-64 en cepa de E.coli	117
Figura 44.	Cromatografía de afinidad de His-SMT3-NMI16-64	118
Figura 45.	Muestras resultantes de la proteólisis de His-SMT3-NMI16-64	119
Figura 46.	Filtración en gel de NMI16-64 tras la eliminación de la cola de histidinas y muestra final utilizada para los ensayos de cristalización	120
Figura 47.	Espectroscopia de Dicroísmo Circular de NMI16-64	121
Figura 48.	Cromatografía de afinidad por histidinas de las proteínas completas BRMS1 y NMI co-expresadas	122
Figura 49.	Proteólisis de la mezcla proteica His-SMT3-BRMS1FL+NMIFL procedente de la primera cromatografía de afinidad y purificación de los productos de digestión	123
Figura 50.	Espectros bidimensionales obtenidos en experimentos de RMN para las proteínas 15N-BRMS151-84 individual y para la mezcla 15N-BRMS151-84 - NMI16-64	125
Figura 51.	Calorimetría de Titulación Isotérmica para caracterizar la interacción de BRMS151-84 y NMI16-64	127
Figura 52.	Cristales obtenidos de la muestra que presenta las proteínas completas BRMS1 y NMI, obtenida mediante co-expresión y co-purificación de las mismas	128
Figura 53.	Alineamiento de la secuencia aminoacídica de los dos dominios de unión a ubiquitina (UBZ1 y UBZ2) de la proteína Tax1BP1 entre diferentes especies	131
Figura 54.	Predicción de estructura secundaria proteica de los dos dominios UBZs de Tax1BP1	132
Figura 55.	Expresión de His-PP-UBZ1, His-PP-UBZ2 y His-PP-UBZ1+2 en cepa de E.coli	133
Figura 56.	Cromatograma y muestras resultantes de la columna de afinidad a histidinas en una purificación de His-PP-UBZ1+2	134
Figura 57.	Muestras resultantes de la proteólisis de His-PP-UBZ1+2 y cromatografía de exclusión por tamaño molecular de la proteína digerida	135
Figura 58.	Gel de poliacrilamida mostrando un elevado grado de pureza de la proteína final UBZ1+2 utilizada para ensayos de cristalización	136

Figura 59.	Secuencia originada tras la expresión en un sistema procariota y purificación del fragmento comprendiendo los aminoácidos 725 al 789 de la proteína Tax1BP1 antes y después de su proteolisis	136
Figura 60.	Cristales de UBZ1+2 obtenidos mediante difusión de vapor	137
Figura 61.	Imagen representativa del patrón de difracción obtenido de un cristal de UBZ1+2	138
Figura 62.	Estrategia utilizada para resolver la estructura de UBZ1+2 de Tax1BP1 mediante SAD	141
Figura 63.	Empaquetamiento en el cristal de las once cadenas de los dos dominios UBZ de Tax1BP1 que conforman la unidad asimétrica	142
Figura 64.	Empaquetamiento de las 11 cadenas de UBZ1+2 en la unidad asimétrica	144
Figura 65.	Diagrama de superficie mostrando la estructura de los dominios UBZ1 y UBZ2	144
Figura 66.	Coordinación entre Cys2-His2 y un átomo de zinc en el motivo “dedo de zinc” UBZ1	145
Figura 67.	Diagrama de Ramachandran para la validación de la estructura de UBZ1+2	146
Figura 68.	Cromatografía de exclusión por tamaño molecular de la mezcla proteica UBZ1+2 - Ubiquitina (Ub)	147
Figura 69.	Cromatografía de exclusión por tamaño molecular de la mezcla proteica formada por UBZ1+2 y por poli-4Ub	149
Figura 70.	Representación esquemática de los dominios de la proteína Itch y construcciones clonadas para su expresión en sistema procariota	150
Figura 71.	Expresión de GST-PP-ItchWW3+WW4, en cepa de E.coli y cromatografía de exclusión por tamaño molecular de la proteína digerida	151
Figura 72.	Cromatografía de exclusión por tamaño molecular de la mezcla proteica conteniendo los dominios UBZ1+2 de Tax1BP1 y WW3+WW4 de Itch	153
Figura 73.	Cristales de los dominios UBZ1+2 en presencia de ubiquitina y de poli-4Ub	154
Figura 75.	Asignación sobre la secuencia de BRMS151-84 del patrón de 7 residuos repetidos en tándem, o “heptada” típico de los motivos estructurados en “coiled-coil”	163
Figura 76.	Superficie de hidrofobicidad e interacciones hidrofóbicas establecidas entre las cadenas A-Asim2 de la estructura de BRM151-84	163
Figura 77.	Discontinuidad en la estructura helicoidal e interacciones polares establecidas en el primer motivo “coiled-coil” de la proteína BRMS1	164
Figura 78.	Posiciones de los residuos y orientación de sus cadenas laterales en las dos hélices antiparalelas A-Asim2 formando el “coiled-coil”	165
Figura 79.	Modelo de estructura elongando 15 residuos en el extremo carboxilo-terminal del motivo “coiled-coil” de BRMS51-84	166
		225

Figura 80.	Asignación del patrón de 7 residuos repetidos en tándem o “heptada”, sobre la secuencia aminoacídica de la construcción clonada para BRMS151-98	167
Figura 81.	Diagrama mostrando las diferencias en los factores de dispersión real e imaginario (f' y f'') del átomo de zinc al usar diferentes longitudes de onda y con ello diferentes energías	170
Figura 82.	Estructura de los dominios “dedos de zinc” de Tax1BP1, mostrando la coordinación con los átomos de zinc del tipo C2H2	171
Figura 83.	Alineamiento de las secuencias consenso entre diferentes tipos de dominios de unión a ubiquitina	173
Figura 84.	Representación de los dominios de unión a ubiquitina de la proteína Tax1BP1, mostrando en un diagrama su estructura secundaria obtenida tras la resolución de la estructura	173
Figura 85.	Superposición de las estructuras tridimensionales de los dominios “dedos de zinc” del tipo C2H2 de unión a ubiquitina UBZ2 de Tax1BP1 y de la proteína polimerasa Eta	174
Figura 86.	Superposición de los dominios “dedos de zinc” del tipo C2H2 y C4 correspondientes a las proteínas Tax1BP1 (UBZ2; en color azul) y Npl4	175
Figura 87.	Representación tridimensional de las estructuras de los dos dominios UBZ de Tax1BP1 en tándem, mostrando las secuencias consenso PPxY que intermediarían la unión con los motivos WW3 y WW4 de la proteína Itch	176
Figura 88.	Representación esquemática de las etapas llevadas a cabo para la determinación de una estructura tridimensional por la técnica de cristalografía de rayos X	186
Figura 89.	Esquema representativo de la deducción de la “Ley de Bragg”	187
Figura 90.	Diagrama de Ramachandran	193

Ecuaciones

Ecuación 1.	Ley de “Lambert-Beer” para la estimación de la concentración de proteína	62
Ecuación 2.	Ley de Bragg	187
Ecuación 3.	Aplicación que responde a la transformada de Fourier	189
Ecuación 4.	Cálculo del factor R, que muestra la adecuación del modelo	192

Abreviaturas

A20	Proteína inducible por factor de necrosis tumoral
ADN	Acido desoxirribonucleico
ARN	Acido ribonucleico
BrMS1	Supresor 1 de Metástasis de cáncer de mama (“Breast cancer Metastasis Suppressor 1”)
C-terminal	extremo Carboxilo-terminal
cAMP	Adenosina monofosfato cíclica
CD	Dicroísmo Circular
cGMP	Guanosina monofosfato cíclica
DTT	Ditiotriol
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	Acido etilendiaminotetraacético
ESRF	“European Synchrotron Radiation Facility”
HDAC	Histona Desacetilasa
IκBα	Inhibidor de NFκB (“IkappaB”)
IKK	Complejo Quinasa IkappaB
IPTG	Isopropil-b-D-tiogalactopiranosido
ITC	Calorimetría de Titulación Isotérmica
N-terminal	Extremo Amino-terminal
NDP52	Proteína Nuclear Dot 52
NEMO	Modulador Esencial del Factor de Transcripción Nuclear kappa B
NFκB	Factor de Transcripción Nuclear kappa B
NLS	Secuencia de Localización Nuclear
NMI	Proteína de Interacción a N-myc (“N-myc Interactor”)
NSCLC	Línea celular de cáncer de pulmón de células no pequeñas
NZF	Dominio “dedo de zinc” de la proteína Npl4
OPN	Osteopontina
p/v	peso/volumen
p50	Subunidad p50 del factor de Transcripción NFκB
PAGE	Electroforesis en gel de poliacrilamida

PAZ	Dominio “Piwi Argonaute Zwiller”
pb	Pares de bases
PCR	Reacción en Cadena de la Polimerasa
PDB	“Protein Data Bank”
PEG	Polietilenglicol
r.m.s.d.	Desviación cuadrática media
r.p.m.	Revoluciones por minuto
RelA/p65	Subunidad RelA/p65 del Factor de Transcripción NFκB
RIP	Proteína de Interacción a Receptor
RMN	Resonancia Magnética Nuclear
RUZ	Dominio de unión a ubiquitina de la proteína Rabex-5
SAB	Seroalbúmina bovina
SAD	“Single-wavelength Anomalous Diffraction”
SDS	Dodecil sulfato sódico
SIRAS	“Single Isomorphous Replacement with Anomalous Scattering”
Ta	Temperatura de anillamiento
Tax	Oncoproteína del virus humano HTLV-1
Tax1BP1	Proteína de Unión a Tax1BP1 (“Tax1 Binding Protein 1”)
Td	Temperatura de fusión o disociación
TRAF	Factor Asociado al Receptor TNF
UBD	Dominio de Unión a Ubiquitina
UBZ	Dominio “dedo de zinc” de unión a ubiquitina
uPA	Activador de Plasminógeno del tipo Uroquinasa
v/v	volumen/volumen